

مطبوعات
مكتبة الملك فهد الوطنية
السلسلة الثانية
(٤٥)

المكتبات الرقمية

تأليف: وليم آرمز

ترجمة
د. جبريل بن حسن العريشي   د. هاشم فرحات سيد

الرياض
١٤٢٧هـ - ٢٠٠٦م

1000110101 01

مقدمة السلسلة

مع مطلع عام ١٩٩٠م، أخذت فكرة المكتبات الرقمية وعملية النشر الإلكتروني تنتقل من اهتمامات غير معلنة لبعض الرؤى المستقبلية إلى أنشطة باتت منافسة للمكتبات التقليدية ولصناعة النشر التقليدي.

وقد كانت العوامل المؤدية إلى ذلك ذات طابع تقني (متمثلة في تقنيات الحاسبات الشخصية، والإنترنت، ونسيج الشبكة العنكبوتية العالمية www) مع أن ثمة مساهمات قد جاءت من جميع فروع المعرفة لتثمن بث المعلومات الجيدة، في الوقت الذي تعد فيه التحديات المصاحبة لإنشاء مكتبات جديدة تحديات اجتماعية واقتصادية وقانونية أكثر منها تحديات تقنية.

وإذا كانت المكتبات الرقمية تقوم الآن على أكتاف أناس يعملون في المكتبات التقليدية، فإن ثمة آخرين من غير المكتبيين أو الناشرين يشاركونهم حمل هذه المهمة. كما هو الحال بالنسبة للشبكة العنكبوتية العالمية التي استحدثت مهناً جديدة، كذلك أصبحت جهود البحث العلمي في مجال المكتبات الرقمية تمثل مجالاً علمياً حيوياً، ينخرط فيه علماء الحاسب الآلي جنباً إلى جنب مع علماء آخرين متخصصين في علوم أخرى كعلماء الاقتصاد والاجتماع، بالإضافة إلى محامين ومكتبيين. وقد أثمر ذلك كله عن نشأة كائن علمي ذي طبيعة متعددة التخصصات وله خبراؤه المختصون.

وقد يبدو متناقضاً أن تؤلف كتب الآن عن المكتبات الرقمية في الوقت

الذي يعتمد فيه معظم الباحثين في هذا المجال في تجميع معلوماتهم الحديثة على الإنترنت بشكل كبير، غير أنهم أنفسهم يدركون تماماً قيمة المطبوعات العلمية التقليدية، كما يدركون أن الابتكارات والبحوث الجارية المبثوثة في مقالات الدوريات، وفي الصحف والمجلات، وأعمال المؤتمرات، قد باتت متاحة مباشرة عبر الإنترنت.

إن هذا الكتاب يتيح فرصة للتأمل، ويعطى تصوراً مستقبلياً لهذا المجال، وهو هدف ينبثق من أهداف سلسلة مطبعة معهد ماساشوسيتس للتقنية MIT عن المكتبات الرقمية والنشر الإلكتروني التي تسعى إلى إتاحة هذه الفرصة، وتوفير هذا المناخ. وآمل أن تكون هذه السلسلة مفيدة للمتخصصين في هذا المجال وغير المتخصصين؛ وسوف يرى المتخصصون كيف ينظر غيرهم من المهتمين بهذا المجال إلى آخر التطورات، وإلى جهود البحث الجارية، وإلى الاتجاهات المستقبلية بعيدة المدى فيه، هذا في الوقت الذي يمكن أن يُقدّم غير المتخصصين إلى المبادرة التي ستؤدي إلى خلق جيل جديد من عالم المكتبات ونمط جديد من صناعة النشر.

وليم آرمز

مقدمة المؤلف

في عام ١٩٩٥م قمت أنا وبعض الزملاء بإنشاء "مجلة المكتبات الرقمية" D-Lib magazine التي تتاح شهرياً على الخط المباشر عبر الإنترنت، والتي أصبحت بمثابة لسان حال مجال البحث في مجال المكتبات الرقمية. ولقد دشنا هذه الصحيفة عن قناعة بأن المكتبات الرقمية مجال يجمع المهتمين من مختلف فروع المعرفة، غير أنهم قلما يعرفون الكثير بعضهم عن بعضٍ، كما كان هدفنا كذلك إنشاء مجلة تعلم الناس بأن ثمة جيشاً من الخبراء يمدون هذا المجال بالمعرفة. دواعي الاهتمام أن علماء الحاسبات الآلية غالباً لا يعون ذلك الفهم العميق للمعلومات التي طورها المكتبيون عبر سنوات عديدة، كما أن المكتبيين والناشرين قد لا يعرفون كذلك أن رواد الإنترنت كانوا يديرون المعلومات المباشرة online informationلعقود مضت؛ وإذا كان المكتبيون والناشرون مدركين بأن مجالاتهم قد تكون مرتعاً للقوى الاقتصادية والقانونية الخارجية، فإن معرفتهم بملابسات ذلك لا تزال محدودة.

لقد أظهرت السنوات القليلة الأولى من عمر مجلة المكتبة الرقمية أن تلك الرؤية كانت ثاقبة، فقد غطت المجلة الكثير من الموضوعات، والتي كانت من بينها: الورشة الأخيرة لقضية وراء البيانات Metadata، ومشروعات التنفيذ التي تبنتها مكتبة الكونجرس، ومبادئ واجهات المستفيدين المطبقة في عمليات البحث والاسترجاع، والمراجعة التاريخية

لمعيار تبادل البيانات الببليوجرافية المعروف بـ "Z39.50 "[١]، والمكتبة الرقمية الخاصة بالحكايات الشعبية اليابانية، وسبل عمل مشروع جستور (JSTOR)للمواد القانونية[٢]، وسبل معالجة مجموعات الرموز البديلة. ومع أن شخصاً واحداً لا يمكن لـه أن يلم بكل تلك المجالات، فإن إدراك مثل تلك الأمور يعد مطلباً لكل من يتطلع أن يكون رائداً في مجال المكتبات الرقمية.

إن هذا الكتاب محاولة لاستعراض جوانب مجال المكتبات الرقمية برمته. وإذا كان للحاسبات الآلية والشبكات أهميتها الأساسية في هذا المجال، فإنها مقومات تقنية لا أكثر، وإن حقيقة المكتبات الرقمية ليست إلا تجسيدًا لجوهر عملية التفاعل القائم بين الناس والمنظمات والتقنيات. كما أن هذا المجال يثير تساؤلات كثيرة عن كيفية استثمار المكتبات ودور النشر لهذه التقنيات الجديدة وكيفية الاستفادة منها ؟ وكيف أن الأفراد يهملون المؤسسات التقليدية [٣] ويبنون مكتباتهم الخاصة بهم؟، وإلى أين ستقود كل تلك الأحداث ؟ ومع أن الإجابة عن هذا التساؤل الأخير قد تبدو سهلة وبسيطة، وهي: لا أحد يدري إلى أين ستقود كل تلك الأحداث! فقد حاولت أن أتفادى التكهن بما سيكون، وأن أركز على وصف الأنشطة والاتجاهات والبحوث

---

(١) Z 39-50 عبارة عن معيار للبروتوكولات التي تكفل القدرة على البحث عن التسجيلات الببليوجرافية واسترجاعها في عدد من نظم التخزين والاسترجاع المتاحة على الخط المباشر (المترجمان).

(٢) JSTOR هو أحد المشروعات التي ترعاها وتمولها مؤسسة أندرو ميلون Andren w. Mellonf . بهدف المحافظة على الأعداد القديمة الهامة من الدوريات القانونية، وتيسير سبل التعامل معها في شكل إلكتروني، بإنتاج صور إلكترونية عالية الجودة منها (المترجمان).

(٣) يقصد المكتبات التقليدية، ودور النشر التقليدية (المترجمان).

الجارية. وبناء عليه فقد حرصت على أن تكون اللوحات التي تحاول كل منها أن تصف جانباً هاماً من جوانب المكتبات الرقمية، أو التقنيات المرتبطة بها، أو توجهات تطبيقها، أو البحث في هذا المجال - في بؤرة اهتمامنا في هذا الكتاب.

إن هذا الكتاب يعكس- بكل تأكيد- خبرتي الخاصة، كما يظهر تحيزي الشخصي. ومن المناسب أن أشير إلى أن اختياري للأمثلة المطروحة في الكتاب قد تركز على ما أعرفه من خلال واقع العمل بالجامعات الأمريكية، مع تقديري ـ بالطبع ـ للجهود الطيبة التي تتم حول العالم، كما أود أن أشير كذلك إلى أنه إذا كان بعض تحيزي قد جاء من غير قصد، فإن البعض الآخر جاء متعمداً بلا شك، كما يمكن الاعتراف بأنني متحيز للمكتبات الرقمية التي تتيح للجميع حرية الوصول إلى المعلومات؛ حيث أثار مشاعري ـ كقارئ ـ أن الإنترنت تتيح لي حرية الوصول إلى رصيد هائل من المعلومات غير المحدودة، ومن ثم فإنني - كمؤلف - أتطلع إلى نشر جهدي البحثي مباشرة عبر الإنترنت ليكون متاحاً للآخرين [1].

---

(١) قام المؤلف بالفعل في عام ٢٠٠٣م بإتاحة نسخة إلكترونية من كتابه هذا مجاناً على الإنترنت .

الفصل الأول

المكتبات والتقنية والناس

إنها لفترة ممتعة في تاريخ المكتبات وصناعة النشر، فلأول مرة يصبح من الممكن تقديم خدمات عالية المستوى؛ حيث يتم اختزان المعلومات في صيغ رقمية، واسترجاعها عبر الشبكات، فبعد أن يتم اختزان المواد في الحاسبات الآلية، فإن الشبكة تقوم بربط هذه الحاسبات أو وصلها بالحاسبات الشخصية الموجودة في مكاتب المستفيدين. وفي عالم المكتبات الرقمية الكاملة لم تعد هناك حاجة البتة إلى المواد الورقية.

ويسعى هذا الكتاب إلى تقديم نظرة عامة على هذا الميدان الجديد، وإن كان يتناول التقنيات بشكل جزئي، فهو يتحدث على نحو متواز عن الناس وعن المؤسسات؛ حيث إن مجال المكتبات الرقمية يضم جوانب متعددة من فروع المعرفة، كما يضم خبراء ذوي خلفيات علمية مختلفة وتوجهات متباينة. ومن ثم فإن هذا الكتاب يسعى إلى وصف الأدوار التي أسهمت بها المجالات المعرفية المختلفة، وكيف تتفاعل فيما بينها، ويناقش كذلك واقع منتجي المعلومات والمستفيدين منها من حيث احتياجاتهم ودوافعهم وحوافزهم الاقتصادية. كما يحلل التغيرات الجذرية التي تحدث الآن في كل من عالمي النشر والمكتبات، ويصف كذلك جهود البحث في مجال التقنيات الحديثة التي يقدم كثير منها عبر الإنترنت، وعلى نسيج الشبكة العنكبوتية العالمية. ومن بين الموضوعات التي تطرح في هذا السياق، الجوانب الفنية للحاسبات الآلية والشبكات، ومهنة المكتبات، وصناعة النشر، فضلاً عن

بعض القضايا الاقتصادية والقانونية.

إن الموضوع الأساسي لهذا الكتاب هو التغير بكل ملابساته الاجتماعية والتنظيمية والقانونية، ونظراً لعدم القدرة على تغطية كل تلك القضايا بشيء من التفصيل في كتاب واحد، فقد طرحت الكثير من البديهيات جانباً حيث تمت معالجتها في أعمال تمهيدية.

وتجدر الإشارة إليه أن معظم الأمثلة التي تطرح في هذا الكتاب مستمدة من البيئة الأمريكية، وخاصة من الجامعات والمجتمع الأكاديمي بشكل عام، مع أن تطور المكتبات الرقمية هو تطور عالمي أسهمت فيه جهات كثيرة، ناهيك عن أن المختصين في الجامعات الأمريكية الكبرى ليسوا وحدهم هم أصحاب الفضل في تطور المكتبات الرقمية، وإن كانوا هم المساهمين الأساسين في ذلك. كما أن هناك الآن فيضاً وتنوعاً كبيراً من الابتكارات في معظم التخصصات، وعلى مستوى جميع دول العالم تقريباً.

إن التعريف غير الرسمي للمكتبة الرقمية هو أنها مجموعة من المعلومات الخاضعة لإدارة جيدة، مع ما يتصل بها من خدمات، حيث يتم اختزان المعلومات في صيغ رقمية، ومن ثم إتاحتها عبر شبكة من الحاسبات. ومن الملاحظ أن الشق الهام في هذا التعريف، هو كلمة المعلومات الخاضعة للإدارة؛ فسيل البيانات الذي يرسل إلى الأرض عبر الأقمار الاصطناعية لا يعد مكتبة، في حين أن تلك البيانات ذاتها يمكن أن تصبح مجموعة لمكتبة رقمية إذا ما تم تنظيمها بشكل منهجي أو مقنن. وهناك كثيرون لا يعدون قاعدة البيانات التي تحتوي على سجلات مالية لإحدى الشركات مكتبة رقمية،

غير أنهم يمكن أن يقبلوا أن تكون مجموعة من تلك البيانات جزءاً من إحدى المكتبات. إن المكتبة الرقمية تحتوي على مجموعات متباينة من المعلومات المتاحة للاستخدام لفئات عديدة من المستفيدين، ومن ثم فهي تتراوح في حجمها من مجرد مكتبة صغيرة جداً إلى مكتبة ضخمة، ويستطيع هؤلاء المستفيدون أن يستخدموا في تعاملهم مع تلك المكتبات أية تجهيزات حاسوبية أو آلية وبرمجيات مناسبة. والقاسم المشترك بين تلك المكتبات هو تنظيم المعلومات واختزانها في الحاسبات الآلية، ثم إتاحتها عبر الشبكات وفق إجراءات معينة بشأن اختيار مواد تلك المجموعات، وتنظيمها، وإتاحتها للمستفيدين، وحفظها لفترات لاحقة.

وإذا كانت المكتبات الرقمية تختلف في كثير من جوانبها عن المكتبات التقليدية، فإنها تتشابه معها بشكل ملحوظ في جوانب أخرى. فالبشر لم يتغيروا بسبب اختراع التقنيات الحديثة، بل سيظلون ينتجون المعلومات التي تحتاج بدورها إلى تنظيم واختزان وتوزيع، كما أنهم سيظلون في حاجة دائمة إلى المعلومات التي ينتجها الآخرون بهدف الإفادة منها لأغراض الدراسة أو البحث أو حتى لأغراض التسلية. وبالرغم من ذلك، فإن الوسائط والطرق التي تستخدم في تنظيمها تتأثر تأثراً كبيراً بالتقنيات، وهذا هو الذي يخلق التغيير، ففي كل عام تتنامى كميات المعلومات المتاحة في الأشكال الرقمية، كما تتحسن التقنيات الداعمة لذلك بشكل مطرد. وهذه التغيرات تخلق في المقابل - وبشكل تراكمي- تغيرات أساسية في طرق إنتاج المعلومات وكيفية استخدامها.

ولكي نعي دور هذه القوى فإن الأمر يتطلب فهماً لأولئك الذين يضطلعون

بمهام تطوير المكتبات؛ فالتقنيات قد أملت طبيعة الخطوة التي يمكن للمكتبات الرقمية أن تكون قادرة عندها على التطور، غير أن الأسلوب الذي تستخدم به التقنيات يعتمد على الناس أنفسهم. وهناك مجتمعان هامان هما مصدرا الكثير من هذه الابتكارات، يتمثل أول هذين المجتمعين في مهنيي المعلومات information professionals، بمن فيهم المكتبيون والناشرون وكثيرون من متعهدي خدمات المعلومات، كخدمات التكشيف والاستخلاص. أما المجتمع الآخر فيشمل الباحثين في علوم الحاسب الآلي ومطوري خدمات الإنترنت. ومن اللافت للنظر أنه وحتى وقت قريب لم يكن ثمة تفاعل جدي بين هاتين الفئتين، بل من المألوف حتى يومنا هذا أن تجد عالماً متخصصاً في الحاسب الآلي لا يعرف شيئاً عن أساسيات مهنة المكتبات، وأن تجد في المقابل متخصصاً في المكتبات لا يعرف عن عملية استرجاع المعلومات إلا المفاهيم القديمة، غير أن بوادر التفاهم بين المتخصصين من كلتا الفئتين قد بدت تتضح مع انطلاقة الويب WEB، وكان من نتيجة ذلك بشكل جزئي أن أصبحت المكتبات الرقمية مجالاً خصباً للبحث، ولكن العامل الأكثر أهمية في ذلك هو الاهتمام المكثف من جانب المستفيدين؛ فالتجهيزات المنخفضة التكاليف والبرمجيات المبسطة أتاحت المعلومات الإلكترونية مباشرة لكل إنسان، ولم يعد المؤلفون في حاجة إلى خدمات الناشرين لتوزيع أعمالهم، كما أن القراء يمكنهم الوصول إلى المعلومات بدون وسيط. والآن هناك الكثير من التطورات المثيرة قدمها أكاديميون أو مهنيون ممن طوروا مكتبات رقمية لاحتياجاتهم الخاصة. ويذكر في هذا الصدد ما لمجال الطب من تقاليد راسخة في التطورات الإبداعية، وما لنظم المعلومات القانونية الرائدة التي طورها القانونيون لخدمة قرنائهم من المحامين، كما أن المتخصصين في مجال

الفيزياء كانوا قد طوروا نسيج الشبكة العنكبوتية العالمية أساساً خدمة لقرنائهم من الفيزيائيين.

جوانب اقتصادية :

لا شك أن التقنية تؤثر في الجوانب الاقتصادية والاجتماعية للمعلومات، والعكس صحيح، فتقنيات المكتبات الرقمية تتنامى بشكل سريع، وتتنامى بالمثل الأطر المالية والتنظيمية والاجتماعية، كما تخلق المجموعات المختلفة التي تعمل على تطوير المكتبات الرقمية أعرافاً اجتماعية مختلفة واتجاهات متباينة نحو الكسب المالي، فدور النشر والمكتبات لهما باع طويل في إدارة المواد التقليدية، وخاصة الكتب والخرائط والصور الفوتوغرافية والتسجيلات الصوتية، وغيرها من المواد الأخرى. وقد طوروا الأطر الاقتصادية والقانونية التي يستند إليها في بيع هذه المواد وشرائها، حيث إن غرائزهم الطبيعية تدعوهم إلى نقل الأفكار التي تخدمهم بشكل جيد عند التعامل مع هذه الأشياء المادية إلى المكتبات الرقمية. أما علماء الحاسبات وغيرهم من المستفيدين العلميين مثل الفيزيائيين، فلهم تقاليد مختلفة، وقد بدأ اهتمامهم بالمعلومات الرقمية منذ أن كانت الحاسبات الآلية غالية الثمن، حيث لم يكن يمتلك الحاسبات العاملة على شبكات المعلومات الأولية سوى الباحثين المدعومين دعماً مالياً كبيراً، وقد كانوا يتبادلون المعلومات بشكل غير رسمي وبلا قيود مع زملائهم وبدون تحمل أية مبالغ مقابل ذلك، وبالرغم من أن الشبكات قد تنامت، فإن مبدأ المعلومات المتاحة بلا قيود open information ظل باقياً، كما إن الإطار الاقتصادي الذي يتم تطويره للمكتبات الرقمية يظهر مزيجًا من اتجاهين، فبعض المكتبات الرقمية تقلد

صناعة النشر التقليدي في طلبها شكلاً من أشكال الدفع قبل أن يتمكن المستفيدون من الوصول إلى مجموعاتها وقبل إفادتهم من خدماتها، في حين أن هناك مكتبات رقمية أخرى تستخدم أسلوباً اقتصادياً آخر، وهو إتاحة الوصول إلى جميع مواردها مجاناً للجميع، وعليه فإن تكاليف إنتاج المعلومات وتوزيعها يقع على عاتق المنتج وليس المستفيد. ويحاول هذا الكتاب تقديم وصف للكثير من نماذج هذين الأسلوبين، كما يحاول كذلك إجراء موازنة بينهما. ومن المؤكد أن لكل منهما مستقبله البعيد الأمد، غير أن طبيعة التوازن النهائي بينهما لا يمكن التنبؤ بها الآن.

لماذا المكتبات الرقمية ؟

يتمثل المبرر الأساس لإنشاء المكتبات الرقمية في الاعتقاد بأنها ستعمل على توصيل المعلومات بشكل أفضل مما كان في الماضي، ومع أن المكتبات التقليدية جزء أساس من المجتمع، فإنها لم تعد في وضعها الذي ينبغي أن تكون عليه. فهل يمكننا عمل الأفضل؟ إن المتحمسين للمكتبات الرقمية يقولون إن الحاسبات الآلية والشبكات قد غيرت بالفعل الأساليب التي يتواصل بها الأفراد بعضهم مع بعضٍ ، وثمة زعم في بعض التخصصات العلمية بأن المهني أو الباحث يؤثر استخدام حاسبه الشخص المتصل بشبكة الاتصالات على ذهابه للمكتبة بحثاً عن المعلومات. وبتوافر المعلومات للمهني على هذا النحو ستصبح متوافرة الآن بشكل مباشر للجميع، فعن طريق الحاسب الشخصي يصبح بمقدور المستفيد تتبع مصادر المعلومات المختزنة في الحاسبات حول العالم. وعلى العكس من ذلك، يدرك الجميع، وخاصة المتحمسين بشدة للمواد المطبوعة، بأنها تشكل جزءاً كبيراً من الحضارة، وأن

دورها الباقي في اختزان المعلومات ونقلها يمكن أن يتغير بشكل تدريجي. ومع أن المعلومات الإلكترونية قد تحل محل بعض الاستخدامات الهامة للأعمال المطبوعة، فليس الجميع يعتبر التحول الكبير نحو المعلومات الإلكترونية أمراً مرغوباً، حتى وإن كان ذلك مجدياً من الناحية الفنية والاقتصادية والقانونية.

وفي الفقرات التالية عرض لبعض الثمرات المتوقعة من المكتبات الرقمية :

المكتبة الرقمية تنقل المكتبة إلى المستخدم :

لكي يستخدم القارئ المكتبة التقليدية عليه أن يذهب إليها، وقد يستغرق ذلك بضع دقائق بالنسبة للقارئ من داخل الجامعة، لكن ليس جميع الناس طلاباً بالجامعة، كما أن المكتبات ليست قريبة من الجميع، يضاف إلى ذلك أن كثيراً من المهندسين والأطباء يجدون صعوبات في الحصول على أحدث المعلومات التي تدخل في إطار اهتماماتهم.

أما المكتبات الرقمية فإنها تنقل المعلومات إلى مكتب المستفيد، سواء أكان في عمله أو منزله، ومن ثم فلم يعد القارئ الذي أصبحت المكتبة الرقمية في متناوله في حاجة لزيارة مبنى المكتبة، فقد أصبحت هناك مكتبة حيثما يكون هنالك حاسب شخصي متصل بإحدى الشبكات.

قوة الحاسب تستخدم لأغراض البحث والتصفح :

قد تبدو الوثائق الورقية مناسبة للقراءة، ولكن الوصول إلى المعلومات المختزنة في تلك الوثائق يمكن أن يكون أمراً صعباً، وبالرغم من آلاف

الأدوات الثانوية والمهارات التي يتمتع بها اختصاصيو المراجع، فإن استخدام المكتبات الكبرى يمكن أن يكون تحدياً قوياً، وجرت العادة أن يدعى أن استخدام المكتبات التقليدية ينمي موهبة اكتشاف الأشياء المفيدة، ذلك لأن رواد هذه المكتبات قد يعثرون على المعلومات عبر مواد قد تبدو قيمتها غير متوقعة لهم، وحقيقة الأمر أن المكتبات مليئة بالمواد المفيدة التي قد لا يكتشفها القراء إلا عن طريق الصدفة فقط.

وفي غالب الأحوال تعد نظم المعلومات المبنية على الحاسبات الآلية بالفعل أفضل من الطرق اليدوية في البحث عن المعلومات، وإن لم تكن بالكفاءة التي يطمح لها أي إنسان، فإنها تعد جيدة، كما أنها تشهد تحسناً، هذا فضلاً عن أن الحاسبات الآلية تكتسب أهمية أو فائدة خاصة في العمل المرجعي حيث يتطلب ذلك تنقلاً متكرراً من مصدر معلومات لآخر.

إمكانية تقاسم المعلومات :

تقتني المكتبات والأرشيفات كثيراً من المعلومات الفريدة، ولا شك أن تحميل المعلومات في صيغ رقمية وإتاحتها على الشبكات يعزز من إتاحتها للجميع، وهناك الآن العديد من المكتبات الرقمية والمطبوعات الإلكترونية يتم حفظها في مواقع مستقلة مركزية، وربما يتم الاحتفاظ بنسخ مكررة قليلة حول العالم. ويعد ذلك تطوراً كبيراً قضى على التكرار المادي المكلّف للمواد قليلة الاستخدام، أو على مشكلة الحصول على المادة الفريدة التي لا يمكن الوصول إليها إلا بالانتقال إلى الموقع الذي تختزن فيه.

سهولة تحديث المعلومات :

لا شك أن كثيراً من المعلومات الهامة تحتاج إلى تحديث مستمر، والمواد المطبوعة يصعب تحديثها، لأن ذلك يعني أن الوثيقة كلها تحتاج إلى إعادة طباعة، وأن تستبعد جميع نسخ الطبعة القديمة وتحل محلها النسخ الجديدة. أما تحديث المعلومات فهو أمر سهل عندما تكون الإصدارة الأصلية في صيغة رقمية ومختزنة في حاسب آلي مركزي.

وتعمل العديد من المكتبات على الاحتفاظ بنسخ متاحة على الخط المباشر من الأدلة والموسوعات وغيرها من الأعمال المرجعية، وبمجرد تلقي النسخ من الناشرين، يتم تحميلها على الحاسب الآلي الخاص بالمكتبة. وخير مثال على ذلك أن لدى مكتبة الكونجرس مجموعة مقتنيات متاحة على الخط المباشر تعرف بمجموعة توماس Thomas، وهذه المجموعة تتضمن آخر مسودات التشريعات التي تعرض على الكونجرس الأمريكي قبل إقرارها.

الإتاحة الدائمة للمعلومات:

إن أبواب المكتبة الرقمية لا توصد أبداً، وهي مفتوحة دائماً على مصراعيها؛ وقد أظهرت دراسة حديثة في إحدى الجامعات البريطانية أن ما يقرب من نصف واقعات الإفادة من المجموعات الرقمية في إحدى المكتبات قد تمت في ساعات إغلاق مبنى المكتبة، يضاف إلى ذلك أن المجموعات المستخدمة لم تعر مطلقاً لقراء خارج حدود المكتبة، ولم يسأ ترتيبها، ولم تسرق، و لم تودع أبداً في مستودع بعيد عن المدينة الجامعية، يضاف إلى ذلك أن مجال المجموعات يتسع إلى خارج حدود المكتبة. كما أن الأوراق

الشخصية المتواجدة في أحد المكاتب، أو في مكتبة معينة في الجانب الآخر من العالم من السهل استخدامها كما تستخدم المواد في المكتبة المحلية.

ومما تجدر الإشارة إليه أن ذلك لا يعني أن المكتبات الرقمية مبرأة من كل عيب، فنظم الحاسبات يمكن أن تتعطل، كما أن الشبكات يمكن أن تكون بطيئة أو لا يعول عليها. ومع ذلك، وإذا ما قورنت بالمكتبة التقليدية، فإن المعلومات في سياق المكتبة الرقمية غالباً ما تكون متاحة متى ما احتاجها المستفيد وأينما شاء.

إمكانية توفير أشكال جديدة من المعلومات :

لم تعد الأساليب الطباعية دائماً هي الوسيلة المثلى لتسجيل المعلومات ونشرها، فقواعد البيانات يمكن أن تكون وسيلة مثلى لتخزين البيانات الخاصة بالإحصاءات الحيوية، ومن ثم يمكن تحليلها بواسطة الحاسبات الآلية واستخراج مؤشرات جديدة منها، كما أن بيانات الأقمار الصناعية يمكن نقلها بطرق مختلفة، وبالمثل فإن مكتبة علم الرياضيات Mathematics Library يمكن أن تختزن المعادلات الرياضية كرموز آلية، ويمكن معالجتها عن طريق بعض البرامج الخاصة بذلك كبرنامج ماثماتيكا Mathmatica، ومابل Maple.

وحتى في حالة تشابه الصيغ، فإن المواد التي تنتج خصيصاً للعالم الرقمي ليست مشابهة تماماً لتلك التي تصمم للنشر في شكل ورقي، فالكلمات المنطوقة لها أثر مختلف عن الكلمات المكتوبة، كما أن المواد النصية المتاحة على الخط المباشر تختلف اختلافاً كبيراً عن الكلمة المنطوقة والمطبوعة، والمؤلف الجيد

هو الذي يغير أسلوبه في الحديث ليتوافق مع الوسيلة التي يكتب فيها، كما أن المستفيدين يجدون سبلاً جديدة للإفادة من المعلومات، والمواد المنتجة أساساً للعالم الرقمي يمكن أن يكون لها حيويتها التي تفتقدها المواد التي تم تحويلها آليًا إلى صيغ رقمية (١)، وهذا يشبه ما يحدث عند عرض أحد الأفلام في حفلة سينمائية، حيث لا يبدو كما هو تماماً عند عرضه على شاشات التلفزة.

إن جميع تلك الثمرات التي أشرنا إليها في الفقرات السابقة نلمسها في واقع المكتبات الرقمية الحالية، غير أن هناك مجموعة أخرى من الثمرات المحتملة التي لم تتجسد معالمها بعد، وإن كان مستقبلها مبشراً، والأمل يحدونا أن تتطور المكتبات الرقمية من مجرد المستودعات التاريخية الساكنة للمواد التي لا تقبل التغيير إلى تقديم الكثير من الخدمات واسعة النطاق التي تفسح المجال لمزيد من التعاون وتبادل الأفكار. ولعل تقنية المكتبات الرقمية تتشابه إلى حد مع التقنية المستخدمة في البريد الإلكتروني والمؤتمرات السمعية عن بعد والتي لم يكن لها من الناحية التاريخية سوى علاقة بسيطة بالمكتبات، مع أن احتمالات التقارب بين هذه المجالات أمر مثير للجدل.

تكلفة المكتبات الرقمية :

إن الثمرة الأخيرة المحتملة للمكتبات الرقمية تتمثل في عملها على توفير الأموال. وعلى الرغم من عدم توافر البيانات القوية عن حجم تكلفة المكتبات الرقمية، فإن هناك بعض الحقائق الأساسية الواضحة في هذا الصدد.

---

(١) عادة يطلق على الفئة الأولى مصطلح المواد الرقمية في العربية و digital materials في الإنجليزية، وتعرف الثانية بالمواد المرقمنة في العربية و digitized or digitalized materials في الإنجليزية (المترجمان).

يتسم واقع المكتبات التقليدية بارتفاع التكلفة، حيث تشغل مباني باهظة الثمن في مواقع حساسة، كما أن المكتبات الكبرى منها كثيراً ما يعمل بها المئات من الموظفين المثقفين - وإن كانت رواتبهم ضعيفة، ومن ناحية أخرى لا تمتلك المكتبات المبالغ المالية الكافية للحصول على كل المواد التي ترغب في الحصول عليها، ولا المبالغ الكافية لإعداد هذه المواد، كما أن عملية النشر هي الأخرى باهظة التكاليف، وتضيف عملية التحول نحو النشر الإلكتروني تكاليف جديدة. ولكي يتمكن الناشرون من تغطية تكاليف تطوير منتجات جديدة فإنهم أحياناً يتحملون دفع مبالغ أكثر، كما هو الحال عند إنتاج طبعة رقمية لنظيرتها الورقية.

وإذا كانت تكلفة المكتبات الرقمية في هذه الأيام أكثر من تكلفة المكتبات التقليدية، فإن أسعار مقومات المكتبات الرقمية تنخفض بشكل سريع. ونتيجة للانخفاض المستمر لتكلفة التقنيات المتصلة بالمكتبات الرقمية، فإنها ستصبح مع مرور الوقت أقل تكلفة، ومن أكثر المجالات التي تشهد انخفاضاً في التكاليف مجالات تخزين المعلومات الرقمية وتوزيعها، ومع أن الانخفاض في التكلفة لن يكون بالدرجة نفسها، فإن بعض الأمور تكون تكلفتها بالفعل أرخص عند معالجتها بالطرق الآلية مقارنة بمعالجتها بالطرق التقليدية. وإن كانت هناك تكاليف أخرى قد لا تشهد انخفاضاً بالمعدل نفسه، بل ربما سترتفع. وبشكل عام يمكن القول بأن هناك فرصة مواتية للحد من تكاليف عمليات النشر وتكاليف المكتبات.

وعلى الرغم من أن التكاليف الأقل انخفاضاً على المدى البعيد لا تعد بالضرورة أخباراً جيدة بالنسبة للمكتبات والناشرين الحاليين، فإن الضغط

لدعم الوسيط التقليدي جنباً إلى جنب مع المجموعات الرقمية الحديثة في المدى القريب عبء ثقيل على الميزانيات. ولأن الناس والمؤسسات يقدرون ثمرات الوصول المباشر للمعلومات والنشر على الخط المباشر، فإنهم مستعدون لصرف المزيد من أموالهم في عمليات التحسيب الآلي والشبكات والمعلومات الرقمية. وعلى أية حال فإن معظم هذه الأموال لا تذهب إلى المكتبات التقليدية، ولكنها تنفق على مجالات حديثة، وبالتحديد الحاسبات والشبكات ومواقع الإنترنت، ومديري مواقع الويب webmasters.

ويواجه الناشرون صعوبات لأن نموذج التسعير الطبيعي لبيع المواد الفردية لا يتناسب وبنية تكاليف النشر الإلكتروني، وإذا كانت معظم تكلفة النشر التقليدي تنصب على إنتاج وتوزيع نسخ الكتب والصور وشرائط الفيديو وغيرها من المواد الأخرى، فإن الأمر يختلف بالنسبة للمعلومات الرقمية. فالتكلفة الثابتة لعملية إنتاج المعلومات وتحميلها على حاسب ما قد تكون ضخمة، مع أن تكلفة استخدامها أو الإفادة منها تكاد لا تذكر نهائياً. ونظراً لأن التكلفة الهامشية قد لا تذكر، فإن كثيراً من المعلومات الموجودة على الشبكات تتاح مجاناً، وبدون أية قيود للوصول إليها. ومع ذلك فليس كل شيء على الشبكات العالمية متاحًا مجاناً، مع أن المتاح كثير جدًا، وهذا من شأنه أن يُقوّض دخل الناشرين. وأن مثل هذه الضغوط تغير بلا شك من القرارات الاقتصادية للمؤلفين والمستفيدين والناشرين والمكتبات.

اللوحة رقم (١ – ١)

رائدان من رواد المكتبات الرقمية

ليست رؤية المكتبة الرقمية رؤية جديدة، وإنما هي مجال تحقق التقدم

فيه بفضل جهود جبارة تبناها كثيرون على مدى فترة زمنية غير قصيرة، ومع ذلك، فهناك قليل من المؤلفين ممن تألقوا بفضل كتاباتهم التي أثارت حماس الأجيال اللاحقة، ويعد كل من فانيفار بوش Vannevar Bush، ج. س. آر. ليكليدر J.C.R. Licklider، من أبرز هؤلاء المؤلفين.

ففي يوليو من عام ١٩٤٥م، نشر بوش ـ الذي كان مديراً للمكتب الأمريكي للبحوث العلمية والتطوير ـ مقالاً في مجلة "Atlantic monthly" بعنوان "كما ينبغي لنا أن نفكرAs we may think"، وقد كان هذا المقال بمثابة عرض موثق للإمكانيات التي يمكن أن تقدمها التقنية للباحثين عند جمعهم للمعلومات وتخزينها والتعرف عليها واسترجاعها. وقد ثبت اليوم صحة الكثير مما تنبأ به منذ خمسين سنة.

وفي سياق مقاله هذا علق بوش قائلاً: "إن أساليبنا في نقل نتائج البحوث ومراجعتها قديمة، ولم تعد الآن بالدقة الكافية لتلبية أهدافها"، كما ناقش التطورات التقنية الحديثة، وإمكانية تطبيقها بشكل معقول في المستقبل القريب، كما قدم ملخصاً لأحد الأساليب الفنية الذي أسماه "ميمكس Memex"، والممتع تاريخياً أن تصميم الميمكس هذا اعتمد على أسلوب استخدم التصوير الفوتوغرافي لاختزان المعلومات، حيث ظل الميكروفيلم ولعدة سنوات هو من أكثر التقنيات المناسبة وغير المكلفة لاختزان المعلومات.

وعلى الرغم من إشارة الكثيرين واستشهادهم ببوش Bush على أنه أول من كشف عن هذه الرؤية الجديدة للمكتبة، فإن ذلك تنقصه الصحة، لأن مقالته اعتمدت على عمل أسبق منها نشر في ألمانيا قبل الحرب العالمية الثانية، ومع ذلك تظل أهمية مقالة بوش هذه باقية لعرضه الجيد للعلاقة بين المعلومات

والبحث العلمي، ولرؤيته الثاقبة للإمكانات الكامنة للتقنية. ومن الجدير بالذكر أن مجلة "Atlantic monthly" قد وضعت نسخة من هذا المقال في موقعها على الإنترنت، ونهيب بكل من له اهتمام بالمكتبات والمعلومات العلمية أن يقرأ هذا المقال.

وفي الستينات من القرن العشرين، كان ليكليدر واحداً من كثيرين بمعهد ماساشوسيتس للتقنية ممن درسوا كيف يمكن للحوسبة الرقمية أن تحدث نقلة في عالم المكتبات. وللاهتمام الكبير الذي كان يوليه ليكليدر ـ كرفيقه بوش ـ بالإنتاج الفكري في مجال العلوم، فقد تكهن بالعديد من التطورات التي ظهرت الآن في عالم الحوسبة الحديثة.

وفي عام ١٩٦٥م، وفي سياق كتابه الموسوم "مكتبات المستقبل Libraries of the future" وصف ليكليدر اتجاهات البحث والتطوير اللازمة لبناء مكتبة رقمية يمكن الإفادة منها على الوجه الصحيح. ومن الملاحظ أن الحوسبة المعتمدة على تقاسم الوقت كانت لا تزال في مختبرات البحث عندما كتب ليكليدر هذا الكلام، وكان البايت الواحد في ذاكرة الحاسب الآلي يكلف دولاراً، ولكنه قدم محاولة جريئة للتنبؤ بما ستكون عليه المكتبة الرقمية بعد ٣٠ عاماً، أي في عام ١٩٩٤م. وقد أثبتت الأيام أن تنبؤاته كانت صائبة في مجملها، مع أنه لم يتنبأ بكل التغيرات التي حدثت في تلك الأعوام الثلاثين. وبشكل عام، يمكن القول إن ليكليدر لم يحسن تقدير تكلفة ما يمكن تحقيقه عن طريق استخدام الأساليب الخارقة، باستخدام كميات كبيرة من قوة الحاسبات الآلية رخيصة الثمن، إلا أنه بالغ في تقدير ما يمكن تحقيقه من تقدم من خلال إمكانات تقنيات الذكاء الاصطناعي، ومن التحسينات التي طرأت على

الأساليب الآلية في معالجة اللغة الطبيعية.

وبالرغم من أن كتاب ليكليدر يعد من أهم الوثائق عن المكتبات الرقمية، فإنه من الكتب التي يصعب الحصول عليها؛ حيث لم يتح على الإنترنت، كما أنه لم يأخذ حقه من الشهرة التي كان ينبغي له أن يحظى بها.

تطورات تقنية :

لقد واجهت المحاولات الجادة الأولى لاختزان معلومات المكتبات في الحاسبات الآلية في أواخر ستينات القرن العشرين مشكلات فنية كبيرة، بما في ذلك التكاليف الباهظة للحاسبات الآلية ذاتها، وواجهات المستفيدين الجيدة، هذا فضلاً عن عدم وجود الشبكات. وحيث إن عملية التخزين كانت باهظة التكاليف، فإن أولى التطبيقات انصبت على المجالات التي كان من الممكن فيها تحقيق عوائد مالية من جراء عملية تخزين مجلدات صغيرة إلى حد ما من البيانات على الخط المباشر. وكان تطوير مكتبة الكونجرس لتسجيلات الفهرسة المقروءة آلياً والمعروفة بمارك MARC من باكورة النجاحات في هذا الصدد.

وقد أثمر استخدام تسجيلات مارك من قبل مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC، بغرض تقاسم عملية إعداد تسجيلات الفهارس بين المكتبات عن تخفيضات هائلة في ميزانيات المكتبات.

ولقد استخدمت خدمات المعلومات المبكرة، مثل خدمات الفهرسة التعاونية، ونظم المعلومات القانونية، وخدمة قاعدة المعلومات الطبية على الخط المباشر المعروفة بمدلاين Medline التي رعتها المكتبة القومية للطب، تلك التقنيات التي كانت متوافرة، حيث كان يتم اختزان كميات قليلة من

المعلومات في حاسبات مركزية ضخمة، وكان على المستفيدين أن يجلسوا أمام منافذ طرفية مخصصة لذلك متصلة بالحاسب المركزي عبر خطوط اتصال بطيئة السرعة، سواء أكانت خطًا هاتفيًا، أو شبكة مخصصة لذلك الغرض(١). وقد كانت هذه النظم تتطلب للتعامل معها مستفيداً مدرباً يقبل العمل من خلال واجهة تعامل معقدة لكنها توفر له وصولاً أسرع لما كان يتم تنفيذه يدوياً، و إلى معلومات لم تكن متاحة محلياً.

إن مثل تلك النظم لم تكن تمثل تهديداً للوثيقة المطبوعة، فكل ما كان يمكن عرضة كان عبارة عن نص غير منسق أو موضوع في صيغة غير مقننة Unformatted، وعادة ما يكون في الخط ذي المسافة الثابتة fixed-space font، وبدون أي إمكانية عرض رسوم تخطيطية أو بيانية أو صيغ رياضية، هذا فضلاً عن عدم توافر الدقة التصويرية المطلوبة لقراءته بشكل يسير. وعندما تضاف جوانب الضعف هذه إلى جوانب الضعف التي تتصل بشاشات الحاسبات المبكرة، المتمثلة في عدم الوضوح الكافي للصور، فضلاً عن انخفاض معدلات نقائها، فإن من المدهش حقاً أن تقتنع الغالبية العظمى من الناس بأن المستفيدين لا يمكن أن يقرؤوا بارتياح شديد من تلك الشاشات.

لقد استطاعت الثلاثون عاماً الماضية أن تتغلب بشكل مطرد على العقبات الفنية التي كانت تواجه مجريات التطور، وفي أوائل التسعينات من القرن العشرين ظهرت سلسلة من التطورات التقنية التي أزالت العقبات الأساسية الأخيرة في بناء المكتبات الرقمية. غير أن بعضاً من التقنيات لا

---

(١) تعرف هذه النظم في الإنتاج الفكري بنظم التجهيز على دفعات خارج الخط المباشر أو بنظم التجهيز على دفعات اختصاراً offline (المترجمان).

يزال في وضع مضطرب، ولكن عمليات استخدام الحاسبات التي تتسم بانخفاض التكاليف قد أثارت انفجاراً في توجه خدمات المعلومات المتاحة على الخط المباشر.

ونشير في الفقرات التالية إلى أربعة مجالات فنية ذات أهمية خاصة بالنسبة للمكتبات الرقمية :

١- الاختزان الإلكتروني أصبح أرخص من الورق :

لا شك أن المكتبات الضخمة مكلفة بصورة كبيرة حتى بالنسبة لأغنى المؤسسات، لأن تكلفة المباني وحدها تستحوذ على ما يقرب من ربع التكلفة الكلية لمعظم المكتبات، وهناك مبانٍ قديمة ضخمة للعديد من المكتبات تفتقر إلى وسائل التحكم البيئي المناسبة، وحتى لو توافرت الأموال الكافية، فإن إمكانات التوسع المكاني غالباً ما يكون من الصعوبة تحقيقها، وخاصة في مراكز المدن المزدحمة أو في داخل المدن الجامعية.

وفي الوقت الذي تتزايد فيه - بشكل مطرد - تكاليف إنشاء مبان حديثة وصيانة المباني القديمة لأغراض اختزان الكتب المطبوعة وغيرها من المواد الأخرى، فإن تكاليف الاختزان الإلكتروني تنخفض بمعدل يصل إلى حوالي ٣٠% سنوياً على الأقل. ففي عام ١٩٨٧م، عندما بدأنا العمل في مكتبة ميركوري الرقمية The Mercury digital library في جامعة كارنيجي ميلون Carnegie Mellon University، طلبنا العديد من الحاسبات الآلية التي كانت السعة التخزينية للواحد منها تصل إلى ١٠ جيجابيت. وكان سعر الحاسب الواحد من هذه الحاسبات يصل إلى حوالي ١٢٠٠٠٠ دولار أمريكي، وفي عام ١٩٩٧م، بلغت تكلفة الحاسب الواحد ذي السعة التخزينية نفسها، بل

بإمكانيات إضافية أفضل، ما يقرب من ٤٠٠٠ دولار فقط. معنى ذلك أن سعر الحاسبات انخفض خلال عشر سنوات إلى ما يقرب من ٩٧%، وهناك ما يدعونا للقول بأن السعر سينخفض في عام ٢٠٠٧م إلى حوالي ٩٧% أخرى.

وفي عام ١٩٨٧م كانت تكلفة اختزان وثائق على الأقراص المدمجة CD-ROM أقل بالفعل من تكاليف اختزان الكتب في المكتبات، واليوم تعد تكاليف اختزان معظم أشكال المعلومات في الحاسبات الآلية أرخص من تخزين الأشكال المادية الأخرى في أي مكتبة تقليدية، وكذلك في العام نفسه، كانت تكلفة التجهيزات تمثل العقبة الكبرى للمكتبات الرقمية، واليوم انخفضت هذه التكاليف بشكل كبير، مع أن اختزان المواد objects الضخمة مثل أجهزة الفيديو الرقمية، أو المجموعة الكبيرة من الصور، أو التسجيلات الصوتية عالية الدقة، لا تزال باهظة التكاليف. وفي غضون عشر سنوات سوف تصبح التجهيزات الغالية جداً قد يصعب شراؤها الآن - منخفضة جداً وبدرجة يندر معها أن يصبح السعر عاملاً مؤثراً في اتخاذ القرارات.

٢ - أجهزة العرض في الحاسبات الشخصية أصبحت مريحة أكثر عند الاستخدام :

لم تعد تكلفة التخزين هي العامل الوحيد (وإذا كانت كذلك فقد كان للمكتبات معاييرها في سبل التصوير المصغر منذ سنوات مضت). وحتى وقت قريب، لم يكن هناك سوى قليل من المستفيدين هم السعداء بالقراءة من شاشة الحاسب الآلي، حيث كانت إمكانات عرض الوثائق على الشاشات ضعيفة جداً، وكان الإجراء المعتاد هو طباعة نسخة ورقية من الوثائق المراد قراءتها. أما الآن فقد شهدت الساحة تطورات هائلة في أجهزة

العرض على الحاسبات، وفي الأحجام التي يتم عرضها على الشاشات، وفي البرمجيات المستخدمة في معالجة المعلومات ونقلها، ورويداً رويدًا بدأ المستفيدون يقرؤون مباشرة من شاشات الحاسبات ـ وخاصة عند عرض المواد التي صممت في الأساس للعرض على الشاشات مثل صفحات الويب. ولا تزال أجهزة العرض الجيدة غالية الثمن إلى حد ما، ولكنها آخذة في الانخفاض عاماً بعد عام، هذا فضلاً عن تحسنها التدريجي. ولكن سيحتاج الأمر إلى وقت طويل حتى تلقى الحاسبات ما تلقاه قراءة الكتب من قناعة، وإن كانت شاشات العرض التي تتمتع بأعلى درجات نقاء الصورة والتي تعمل عليها مختبرات الأبحاث الآن تبدو غاية في الجمال.

٣ – الشبكات عالية السرعة أصبحت أكثر انتشاراً:

لقد بات نمو شبكة الإنترنت خلال تسعينات القرن العشرين ظاهرة ملموسة؛ حيث تتنافس شركات الاتصالات بعيدة المدى لتقدم خدمات الإنترنت المحلية والبعيدة عبر الولايات المتحدة، كما وصلت خطوط الربط الدولية Links إلى كل قطر في العالم تقريباً، وباتت لكل شركة ضخمة شبكاتها الداخلية، وأنشأت الجامعات الشبكات المحلية داخل تجمعاتها، كما يستطيع الأفراد شراء حق الحصول على خدمات الاتصال Dial Up منخفضة التكاليف من منازلهم. ومع ذلك فلا تزال هنالك الكثير من جوانب الضعف، حتى في الولايات المتحدة الأمريكية، بل هناك بعض الدول التي لم يتوافر لها حتى الآن إمكانات الاتصال بالإنترنت نهائياً. وبالرغم من ذلك فإن تلقي المعلومات من الإنترنت في العديد من الدول يعد أكثر سهولة من

الحصول على الكتب والدوريات المطبوعة بالطرق التقليدية القديمة.

٤ – الحاسبات الآلية أصبحت قابلة للنقل :

بالرغم من أن المكتبات الرقمية تعتمد في عملها على الشبكات، فإن إمكاناتها تعززت بشكل كبير بفضل تطور الحاسبات المحمولة Laptop، حيث يمكن للمستفيد من خلال ربط الحاسب المحمول بوصلة الشبكة أن يجمع ما بين مصادر المكتبة الرقمية على الإنترنت مع عمله الشخصي المحمل على الحاسب، وعندما يقوم المستفيد بفصل حاسبه الشخصي عن الاتصال بالإنترنت، فإن نسخاً من مواد المكتبة التي اختارها يمكن أن يتم الاحتفاظ بها للاستخدام الشخصي. وقد تطورت في السنوات القليلة الماضية، إمكانات الحاسبات الآلية المحمولة، وتحسنت كذلك نوعية شاشاتها بشكل كبير، ناهيك عن الانخفاض المستمر في أسعارها.

إمكانات الوصول إلى المكتبات الرقمية:

جرت العادة في عالم المكتبات التقليدية أن تكون متاحة للمستفيدين الذين ينتسبون للمؤسسة التي تتواجد بها تلك المكتبات أو التي ترعاها، ففي الولايات المتحدة، تمتلك الجامعات وبعض المؤسسات الأخرى مكتبات رائعة، غير أن الغالبية العظمى من الناس ليسوا من منسوبي تلك المؤسسات. وإذا كانت مكتبة الكونجرس تتيح كثيراً من مجموعاتها وتفتح أبوابها لجميع من تجاوزوا الثامنة عشرة من أعمارهم، فقليل من المدن هي التي تتوافر بها مكتبات عامة جيدة، ومن ثم فمعظم الناس مقيدون بالمجموعات الصغيرة التي تقتنيها مكتباتهم العامة المحلية، بل إن العلماء نادراً ما تتوافر لهم تسهيلات

مكتبية كافية، كما أن الأطباء في المراكز الطبية الكبرى تتوافر لهم مكتبات جيدة، أما من يعملون في مراكز طبية بأماكن بعيدة فقلما تتوافر لهم تلك المكتبات. ولقد كان عدم توافر إمكانية الوصول للمكتبة الهندسية أحد الدوافع التي أثارت اهتمام معهد المهندسين الكهربائيين والإلكترونيين Institute of Electrical and Electronics Engineers (IEEE) إلى المبادرة بعملية النشر الإلكتروني.

ولأن استخدام المكتبة الرقمية أو الإفادة منها يتطلب أن يمتلك المستفيد حاسباً آلياً متصلاً بالإنترنت، فإن العديد من المؤسسات في الولايات المتحدة تقدم لموظفيها حاسبات شخصية، كما لجأت إلى ذلك بعض المؤسسات الأخرى منذ سنوات مضت، وهناك على المستوى المحلي برامج لتزويد المدارس والمكتبات العامة بالحاسبات الآلية. أما الأشخاص الذين كان عليهم توفير حاسبات لأنفسهم، فقد كان الاتصال بالإنترنت يكلفهم ما لا يقل عن ٢٠٠٠ دولار كقيمة للتجهيزات، وربما يصل إلى ٢٠ دولاراً في الشهر للاتصال بشبكة الإنترنت من خلال شبكة الهاتف العمومية وبالقدر اليسير من مهارة التعامل مع الإنترنت. وبمبلغ مالي أكثر من ذلك بقليل يمكن الحصول على حاسب آلي (ذي مواصفات عالية، والتمكن من إجراء اتصال جيد عالي السرعة بالإنترنت. وإذا كانت هذه تعد استثمارات صغيرة بالنسبة للمهني متيسر الحال، فإنها يمكن أن تمثل عقبة للآخرين، فقد قدر أن ٩٥% من سكان الولايات المتحدة في عام ١٩٩٨م يقطنون في أماكن يسهل الاتصال فيها بالإنترنت بشكل معقول، وسوف تزداد هذه النسبة بشكل كبير مع مرور

الأيام.

أما خارج الولايات المتحدة فالموقف يختلف، إذ إن قليلاً من الدول هي التي تتوافر لها الخدمات المكتبية بالقدر نفسه من الجودة التي تتوافر في الولايات المتحدة، فعلى سبيل المثال، يقال بأن عملية التوزيع الآمن للدوريات العلمية في جامعات المكسيك يعد ضرباً من المستحيل، حتى مع توافر الموارد المالية الكافية، كما أن الحاسبات الآلية والشبكات متوافرة بشكل جيد في بعض الدول، ولكن تكاليف التجهيزات مرتفعة في معظم الحالات أكثر مما هي عليه في الولايات المتحدة، ومن ثم فإن القوة الشرائية منخفضة، كما أن هناك قدراً كبيراً من الاحتكار لعمليات الاتصالات مما يزيد من تكاليفها بشكل كبير، ناهيك عن عدم توافر الدعم الكافي للبنية التحتية اللازمة في تلك الدول. وإذا كانت المكتبات الرقمية تعمل على توفير المعلومات للكثيرين ممن يفتقدون المكتبات التقليدية، فليس من السهل توفير الوصول إلى الإنترنت لكل مكان.

إن العامل الذي يجب أن يؤخذ بعين الاعتبار عند التخطيط للمكتبات الرقمية هو التفاوت الكبير في مدى جودة التقنيات المتوافرة للمستفيد. فقليل من المحظوظين هم ممن تتوافر بمكاتبهم أحدث أجهزة الحاسبات الآلية المتصلة بشبكات اتصال عالية السرعة بالإنترنت، وبأحدث البرمجيات، ويتوافر لهم من ناحية أخرى فنيون مهرة قادرون على تهيئة التجهيزات وضبطها وحل المشكلات، وتحديث البرمجيات. أما الغالبية العظمى من الناس من ناحية أخرى، فلا يتوافر لهم سوى الفتات، فأجهزتهم ربما تكون عتيقة، وبرمجياتها عقيمة، والاتصالات بالإنترنت معوقة، وأعضاء فريق الدعم الفني قليلو الخبرة ومثقلون بالمهام. وإن من أهم التحديات التي تواجه

تطوير المكتبات الرقمية بناء نظم تستفيد من مزايا التقنيات الحديثة، ولكنها تستطيع في الوقت نفسه التعامل بأداء جيد في الظروف الأقل مثالية.

مفاهيم ومصطلحات أساسية :

كثيراً ما تكون المصطلحات عائقاً عند مناقشة قضايا المكتبات الرقمية؛ فالذين بنوا المكتبات الرقمية ينتمون لتخصصات متباينة، وهم قد جلبوا معهم عند تعاملهم مع المكتبات الرقمية مصطلحات تلك التخصصات، وكثيرا ما تكون بعض هذه المصطلحات ذات معان اصطلاحية أو مفاهيم ضمنية اجتماعية أو مهنية أو قانونية أو فنية يمكن أن تعوق الحوار الذي يدور بين أناس لهم تلك الخلفيات المختلفة. وثمة كلمات بسيطة لكنها تعني أشياء مختلفة عند أناس مختلفين. فعلى سبيل المثال، كلمتا "ينسخ Copy" و "ينشر publish" تشيران إلى معان مختلفة عند كل من المتخصصين في مجال الحاسبات وعند الناشرين والمحامين. كما أن الاستخدام الإنجليزي الشائع لا يعني الشيء نفسه بالنسبة للمهنيين، وقد تتفاوت لهجات اللغة الإنجليزية المتداولة حول العالم تفاوتاً كبيراً فيما بينها من مكان لآخر، يضاف إلى ذلك أن المناقشات حول المكتبات الرقمية لم تعد حكراً على اللغة الإنجليزية وحدها.

وثمة بعض الكلمات التي تسبب شيئًا من الخلط أو سوء الفهم بدرجة تدفع إلى عدم تداولها في المناقشات حول المكتبات الرقمية؛ فبالإضافة إلى كلمتي "ينسخ"، و"ينشر" اللتين سبقت الإشارة إليهما، هناك قائمة مطولة بكلمات أخرى مثل: "وثيقة" Document، و"كائن أو عنصر Object" و"عمل Work". وهي من الكلمات التي يجب استخدامها بحذر شديد، كما أن معانيها الدقيقة يجب أن توضح بدقة عند تداولها، وهناك بعض الكلمات الأخرى

على الرغم من عدم غموضها الكبير، إلا أن معانيها تختلف اختلافاً كاملاً باختلاف السياقات التي ترد فيها، فعلى سبيل المثال، يجب في سياقات محددة التفرقة بين كلمة "الصورة الفوتوغرافية Photograph" (التي تعني صورة ما مطبوعة أو مرسومة على الورق)، وكلمة "الصورة المرقمنة Digitized Photograph" (التي تعني مجموعة من البتات Bits المختزنة في الحاسب الآلي)، ومع ذلك ففي معظم الأحيان، لا تكون مثل هذه الدقة إلا مجرد حذلقة لا مبرر لها.

وأحاول في هذا الكتاب أن أسير على نهج غالبية المتمرسين في مجال المكتبات الرقمية، وأن استخدم المصطلحات بشكل غير رسمي حيثما يكون المعنى واضحاً من السياق.

المجموعات Collections :

للمكتبات الرقمية أن تحتفظ بأية معلومات يمكن ترميزها كسلسلة متتابعة من البتات Sequences of bits، وفي بعض الأحيان تكون هذه المعلومات عبارة عن نسخ مرقمنة لوسائط تقليدية، مثل: النصوص والصور والموسيقى، والتسجيلات الصوتية والأوصاف والتصميمات وغيرها الكثير. وحيث إن مجال المكتبات الرقمية مجال متسع، فإن محتوياتها غالباً ما تقتصر على النظائر الرقمية للمواد التقليدية أو المادية الأخرى، ولكن أكثر من ذلك حيث يمكن أن تشتمل على مواد ليس لها نظائرها التقليدية كالبيانات الناتجة عن التجارب العلمية، وبرامج الحاسب، وألعاب الفيديو، ومراصد البيانات.

البيانات Data وما وراء البيانات Metadata [١] :

يمكن أن تقسم المعلومات التي يتم اختزانها في المكتبة الرقمية إلى فئتين: البيانات، وما وراء البيانات. والبيانات مصطلح عام يستخدم ليصف المعلومات التي يتم ترميزها في صورة رقمية، وسواء استخدمت هذه الكلمة في صيغة المفرد أو في صيغة الجمع فهي مصدر للجدل، وقد عاملتها في هذا الكتاب، وعلى ما جرت عليه العادة في مجال الحاسبات الآلية على أنها من الأسماء المفردة التي تستخدم للتعبير عن اسم الجمع Singular Collective Noun).

أما ما وراء البيانات فهي البيانات التي تصف البيانات. ومع أن كثيرين لا يستسيغون هذه الكلمة، إلا إنها أصبحت شائعة التداول، وهناك عدة أنواع من "ما وراء البيانات"، هي: ما وراء البيانات الوصفية Descriptive metadata (مثل المعلومات الببليوجرافية)، وما وراء البيانات التركيبية أو البنائية structrcal metadata (وهي المعلومات عن الصيغ Fromats والبنى أو التراكيب Structures)، وما وراء البيانات الإدارية Administrative metedata (التي تشمل معلومات عن الحقوق وتصاريح الدخول والمعلومات الأخرى المستخدمة لأغراض إدارة الوصول إلى المعلومات)، وأحد عناصر "ما وراء البيانات" هو "المحدد" Identifier ـ الذي يعّرف أو يحدد عملاً أو عنواناً معيناً إلى العالم الخارجي.

وغالباً ما تعتمد التفرقة بين مصطلحي "البيانات" و "ما وراء البيانات"

---

(١) هناك اجتهادات عربية كثيرة لترجمة لهذه الكلمة منها بيانات البيانات، وما وراء البيانات، والبيانات المساندة، الميتاديتا (المترجمان).

على طبيعة السياق، فالبطاقات في الفهارس وأدوات الاستخلاص تعامل عادة على أنها "ما وراء بيانات"؛ لأنها تصف بيانات أخرى، لكنها تعامل في وقت آخر على إنها "بيانات"، وخاصة في الفهرس المباشر online cataloge، أو في مرصد بيانات للمستخلصات.

المواد items في المكتبة الرقمية:

لم يوضع بعد مصطلح عام للتعبير عن الوثائق المختزنة في المكتبة الرقمية، وقد استخدمنا في هذا الكتاب عدة مصطلحات، وأشهر هذه المصطلحات العامة المستخدمة هو مصطلح "مادة Material"، الذي يشير إلى أي شيء يمكن أن يختزن في المكتبة، أما كلمة "وحدة أو مادة Item" فهي مرادفة بشكل أساس لكلمة "مادة Material"، وإن كانت أي من هاتين الكلمتين لا تتضمن ما يشير إلى طبيعة مضمون المعلومات أو شكلها أو وجهة نظر المستفيد فيها.

وكلمة " مادة Material" يمكن أن تستخدم لوصف العناصر المادية أو المعلومات المتواجدة في صيغ رقمية Formats. ومصطلح مادة رقمية Digital Material يستخدم عادة عندما تكون صفة "الرقمية" ضرورية لتأكيد طبيعة هذه المادة. أما المصطلح الأكثر دقة وهو "العناصر أو الكائنات الرقمية" Digital Objects فيستخدم لوصف مواد معينة مختزنة في المكتبة الرقمية، وبالطبع فإن الكائن الرقمي Digital objects يتألف من البيانات، ومن ما وراء البيانات المتصلة بها، ومن المحدد Identifier.

ومع أن هناك بعض الناس يحلو لهم وصف أي مادة في المكتبة الرقمية بأنها "وثيقة Document". فإن هذا الكتاب يتحفظ على استخدام هذا

المصطلح لأي من النص المرقمن digitized text، والعنصر أو الكائن الرقمي، اللذين تكون بياناتهما بمثابة النظير الرقمي للوثيقة التقليدية.

الكائنات أو العناصر المكتبية Objects Library :

مصطلح الكائن أو العنصر المكتبي من المصطلحات المفيدة لأنه يعبر عن وجهة نظر المستفيد فيما هو مختزن في إحدى المكتبات، فمن المحتمل مثلاً أن تختزن إحدى مقالات مجلة إلكترونية في حاسب معين كعناصر منفصلة متعددة (وهي صفحات النص المرقمن، والرسوم البيانية، وربما حتى برامج الحاسب أو المواد التي تم ربطها معها والمختزنة في حاسبات بعيدة)، ومن وجهة نظر المستفيد فإن هذا يعد كائناً مكتبياً واحداً مكوناً من عناصر رقمية متعددة.

وللكائنات أو العناصر المكتبية بناؤها الداخلي، فهي عادة ما تشتمل على كل من "البيانات" و ما يرتبط معها من ما وراء البيانات. على أن تستخدم ما وراء البيانات البنائية لوصف الصيغ وبيان طبيعة العلاقة بين أجزائها.

أساليب عرض الكائن أو العنصر الرقمي، ووسائل توزيعه أو بثه، والشكل المختزن فيه :

إن الشكل الذي تختزن فيه المعلومات في المكتبة الرقمية قد يكون مختلفاً تماماً عن الشكل الذي تستخدم به، فالمحاكي الذي يستخدم في تدريب الطيارين يمكن أن يختزن كبرامج آلية متعددة، وتركيبة بيانات، وصور مرقمنة، وبيانات أخرى، وهذه تعد هي الأشكال التي يختزن فيها هذا الكائن، أما المستفيد فيراه كسلسلة من الصور والأصوات المركبة في تتابعات

منضبطة.

ويستخدم بعض الناس مصطلح "عرض أو تقديم presentation" للإشارة إلـى مـا يعـرض للمستفيد، وقد يكون هذا المصطلح مناسباً، لكن هناك مصطلحاً آخر أكثر عمومية وهو "البث أو التوزيع" الذي يؤكد على أن عملية النقل أو الإرسال من الشكل المختزن إلى المستفيد تتطلب تنفيذاً لبعض برامج الحاسب.

وعند تلقي المعلومات الرقمية من قبل الحاسب الآلي للمستفيد، يقوم بتحويلها إلى الشكل الذي تقدم به للمستفيد كما لو كانت معروضة على شاشة الحاسب، وربما يصاحبها مسارات صوتية أو بعض مظاهر العرض الأخرى. وعادة ما تعرف عملية التحويل هذه بالنقل أو الترجمة rending.

الأعمال Works والمحتوى Content :

تتسم عملية إيجاد المصطلحات التي تصف المحتوى بالصعوبة البالغة. ذلك أن اللغة الإنجليزية مرنة جداً، وأن معاني الكلمات تختلف باختلاف السياق. فتأمل على سبيل المثال أغنية "هدايا متواضعة Simple Gifts"، حيث يمكن لهذه العبارة - اعتماداً على السياق - أن تشير إلى الأغنية كعمل ما مصحوب بالكلمات والألحان، أو إلى معدل الأغنية أو ترتيبها، أو إلى أداء الشخص الذي يغنيها، أو إلى تسجيل الأداء، أو إلى نسخة من الموسيقى محملة على قرص مدمج، أو إلى قرص مدمج محدد، أو إلى عملية سماع الموسيقى من التسجيل، أو إلى الأداء الذي تم ترميزه في مكتبة رقمية، أو إلى مجالات أخرى إضافية للأغنية. وإن مثل هذه الفروق هامة بالنسبة لصناعة الموسيقى لأنها تحدد الطرف الذي يتلقى المبالغ المالية التي تدفع نظير الأداء الموسيقي أو تسجيل الأغنية.

وقد حاول العديد من الباحثين في مجال المكتبات الرقمية أن يضعوا بناءً هرمياً للمصطلحات التي يمكن تطبيقها لكل الأعمال والعناصر المكتبية، وهو هدف قوي وله أهميته، لكنه محفوف بالمخاطر؛ فالمواد المكتبية تتفاوت بشكل كبير لدرجة أن تصنيف بعض الأنواع يمكن أن يتم بشكل جيد، لكنه يفشل في وصف مواد أخرى بالقدر نفسه من الدقة.

ومع أن كلتا الكلمتين "عمل" و "محتوى" لهما القدر ذاته من الأهمية، فإن معظم الناس يستخدمون كلمة "المحتوى" بدون تحديد، ويسير هذا الكتاب على هذا المنوال، فهذه الكلمة تستخدم في أي سياق عندما يكون التركيز على المواد المكتبية التي ليست في شكل بتات أو بايتات يتم معالجتها عن طريق الحاسب، ولكن كمعلومات لها أهميتها لمستفيد ما، فنحن لا يمكن أن نعرف أو نحدد المحتوى أو نحدده، ولكننا نعرفه عندما نراه.

وبينما تستخدم كلمة "المحتوى" بغير تحديد كمصطلح عام، فإن كلمة "العمل" تستخدم بشكل أكثر تحديداً، وهناك تعريف دقيق للعمل الأدبي Literary work ورد في سياق قانون حق المؤلف الأمريكي، يقول إنه محتوى مجرد (كمنظومة من الكلمات أو النوت الموسيقية) غير مقيد بأي عملية تمثيل مختزن أو عرض أو أداء محدد. وغالباً ما يستخدم هذا الكتاب كلمة "عمل work" بهذا المعنى تقريباً، مع أن ذلك ليس دائماً بالدقة الكافية.

الأشخاص المرتبطون بالمكتبات الرقمية :

هناك عدة كلمات تستخدم لوصف الأفراد المرتبطين بالمكتبات الرقمية. فمنتجو المعلومات في عرف المكتبة هم المؤلفون والمؤلفون الموسيقيون والمصورون ورسامو الخرائط والمصممون، وأي شخص آخر ينتج عملاً

فكرياً، ومن بين هؤلاء من يكون محترفاً أو مهنياً ومنهم من يكون هاوياً، كما أن بعضهم يعمل بشكل فردي، وبعضهم الآخر ينخرط في أعمال الفريق أو الأعمال المشتركة. ولكل منهم دوافعه المتباينة لإنتاج المعلومات.

وثمة مجموعة أخرى تتألف من المستفيدين من المكتبة الرقمية. واعتماداً على السياق، فإن هؤلاء المستفيدين يمكن وصفهم بمصطلحات مختلفة. ففي المكتبات غالباً ما يعرفون بالقراء readers أو الزبائن أو الرواد patrons، وإن كانوا في بعض الأحيان يعرفون بالجمهور audience أو العملاء customer. وإذا كان ذلكم هو الحال بالنسبة للمكتبات التقليدية، فإن الموقف قد يختلف في المكتبات الرقمية؛ حيث يكون منشؤوها أنفسهم المستفيدين منها. وفي المجتمع الأكاديمي يستفيد الباحثون والعلماء من المكتبات باعتبارها مصادر لبحوثهم، ثم يقومون بعد ذلك بنشر نتائج بحوثهم في أشكال تصبح جزءاً من مجموعات المكتبة.

وهناك مجموعة أخيرة مرتبطة بالمكتبات الرقمية، وهي مجموعة يتسع مجال تغطيتها حيث تضم كلاً من يدعم المنشئين والمستفيدين، وعليه فهي تضم المتخصصين في الحاسب، وأمناء المكتبات، والناشرين، والمحررين، وآخرين كثر. وجميعهم يمكن أن يطلق عليهم مديرو المعلومات Information manager، وقد أوجدت بيئة الشبكة العنكبوتية العالمية وظيفة جديدة، هي مدير موقع الويب webmaster، وهو شخص واحد يمكن أن يكون منشئاً ومستفيداً أو مديراً للمعلومات.

الحاسبات الآلية والشبكات :

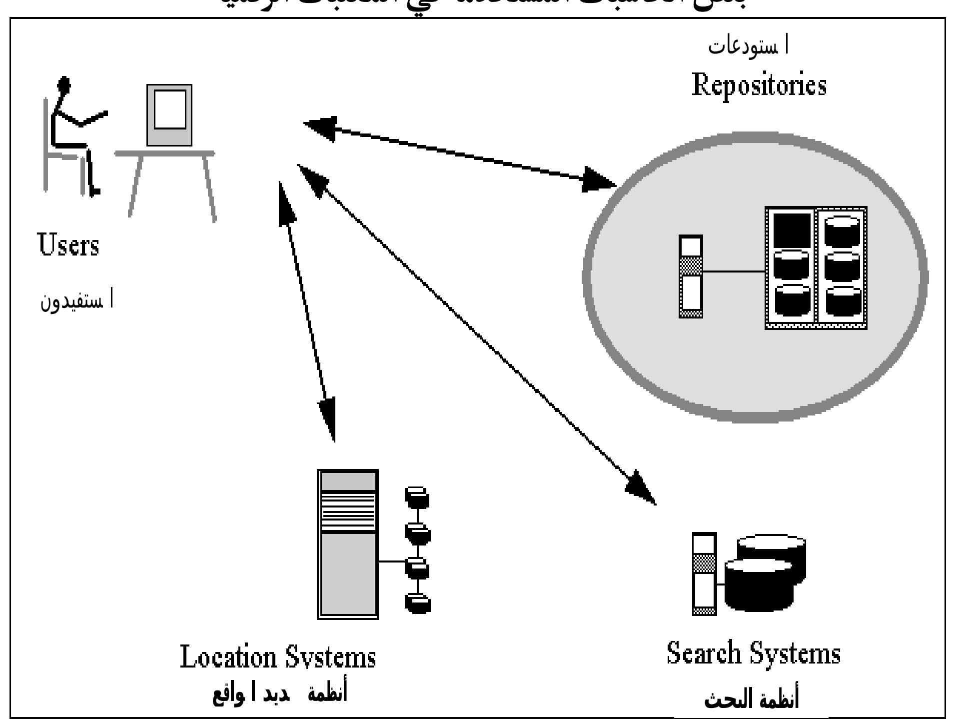
تتكون المكتبة الرقمية من مجموعة من الحاسبات الآلية المترابطة عن

طريق شبكة اتصالات، وتعد الإنترنت هي الشبكة الغالبة في هذا السياق، ولعل تميزها بالمرونة وانخفاض تكلفتها وسعة انتشارها على مستوى العالم هي العوامل الأساس في تطور المكتبات الرقمية.

ويصور الشكل رقم (١-١) بعضاً من الحاسبات التي تستخدم في المكتبات الرقمية. ولهذه الحاسبات ثلاث وظائف تتمثل في مساعدة المستفيدين على التعامل مع المكتبة، والقيام بمهام اختزان مجموعات المواد، وتقديم الخدمات.

الشكل رقم (١-١)

بعض الحاسبات المستخدمة في المكتبات الرقمية



وفي سياق المصطلحات المستخدمة في مجال الحاسبات الآلية، يعرف كل شخص يتعامل مع الحاسب بأنه "مستخدم حاسب" أو ببساطة "مستخدم".

وهذا المصطلح الواسع يشمل كلاً من المنشئين أو المنتجين، والمستفيدين من المكتبات، ومهنيي المعلومات، وأي شخص آخر يتعامل مع الحاسب الآلي. أما الحاسب الآلي الذي يستخدم للوصول إلى المكتبة الرقمية فيعرف "بالعميل Client". وفي بعض الأحيان تتفاعل الحاسبات العميلة هذه مع المكتبة الرقمية بعيداً عن تدخل المستفيد أو العنصر البشري. ومن أمثلة الحاسبات العميلة التي تقوم بذلك "أجهزة الروبوت Robots" التي تقوم بفهرسة مجموعات المكتبة آلياً، و"أجهزة الإحساس sensors" التي تقوم بتجميع البيانات عن الأحوال الجوية وتزويد المكتبة الرقمية بها.

وهناك كذلك المستودعات Repositories وهي الحاسبات التي تقوم باختزان مجموعات المعلومات وتوفير إمكانية الوصول إليها، كما أن هناك "الأرشيف Archive" وهو مستودع منظم يستخدم لأغراض حفظ المواد لفترات طويلة.

هناك نوعان من الخدمات التي تقدمها المكتبات الرقمية، تعرف الأولى بالخدمات البحثية، وتعرف الثانية بخدمات تحديد المواقع. وتتمثل الخدمات البحثية في توفير الفهارس والكشافات وغيرها من الأدوات التي تعمل على مساعدة المستفيدين على إيجاد المعلومات، أما خدمات تحديد المواقع فتتمثل في التعرف إلى مواقع وجود المعلومات وتحديد أماكن هذه المواقع.

ويحدث في بعض الأحيان أن تتواجد حاسبات أخرى بين الحاسبات العميلة من جهة والحاسبات التي تختزن المعلومات من جهة أخرى، وتعرف

هذه بحاسبات المخازن البديلة [١] Mirrors ، وحاسبات الذاكرة المخبأة Caches التي تقوم باختزان نسخ مكررة من المعلومات لأغراض تسريع الأداء وتوفير ضمانات الثقة. وينصب عمل حاسبات المخازن البديلة على عمل نسخ أخرى من أكبر قدر من المعلومات، في حين يتركز عمل حاسبات الذاكرة المخبأة على اختزان آخر المعلومات المستخدمة فقط. وبالإضافة إلى هذين النوعين هناك الملقمات Proxies، والبوابات Gateways التي تعمل على توفير جسور فاصلة بين الأنواع المختلفة من نظم الحاسبات، وهي ذات أهمية خاصة في التوفيق بين النظم ذات المواصفات الفنية المتعارضة.

أما المصطلح العام " الخادم [٢] Server" فيستخدم لوصف أي حاسب آخر غير الحاسب العميل، ويمكن لجهاز خادم واحد أن يقوم بالعديد من الوظائف التي أشرنا إليها من قبل، وربما يعمل مستودعًا، ومقدم خدمة بحث، ومقدم خدمة تحديد مواقع، وعلى العكس من ذلك، يمكن أن توزع الوظائف الفردية على أكثر من جهاز خادم واحد، فعلى سبيل المثال يعد نظام تسمية المجال Domain-Name System، وهي خدمة تحديد الأماكن على الإنترنت، خدمة فردية لكنها متكاملة تعمل على الآلاف من الحاسبات الخادمة الفردية.

أما النظام الموزع أو اللامركزي Distributed system، فهو مجموعة من الحاسبات التي تعمل كفريق لتقديم الخدمات للمستفيدين. وتعد المكتبات الرقمية واحدة من بين أكثر النظم اللامركزية المعقدة والطموحة التي بنيت حتى الآن.

---

(١) يمكن تسميتها بالمطابقة أو العاكسة.

(٢) لعل الكلمة العربية "النادل" هي الأوقع في هذا السياق، غير أن كلمة "الخادم" أصبحت من الكلمات المتداولة على نحو واضح لذلك رأينا استخدامها (المترجمان) .

كما أن الحاسبات الشخصية للمستفيدين والتي تقوم بمهمة تبادل الرسائل مع الحاسبات الخادمة، هي أنواع معروفة جداً، وتديرها آلاف من المؤسسات المختلفة، كما إنها تشغل البرمجيات التي تتراوح ما بين أحدث ما توصلت إليه التقنية والأنواع العتيقة. أما مصطلح "التشغيل المتداخل أو البيني Interoperability" فيشير إلى مهمة بناء خدمات مترابطة للمستفيدين عندما تكون المكونات الفردية مختلفة من الناحية الفنية، وتديرها مؤسسات متباينة. وهناك من يزعم بأن جميع المشكلات الفنية التي تواجه المكتبات الرقمية هي بسبب عملية التشغيل المتداخل، وقد يكون في ذلك مغالاة، ولكن من المؤكد أن التشغيل المتداخل يعد تحدياً أساسياً بالنسبة لجميع مجالات المكتبات الرقمية.

تحدي التغيير :

إذا كانت التقنية الرقمية مثيرة إلى هذا الحد، فما الذي يمنع أي مكتبة تقليدية أن تتحول مباشرة وبشكل كامل إلى مكتبة رقمية؟ وإذا كان طرف من الإجابة عن هذا التساؤل يتمثل في أن تقنية المكتبات الرقمية لا تزال غير ناضجة، فإن الطرف الآخر يتمثل في أن التحدي شيء أكثر من التقنية. وأن التحدي الحقيقي يكمن في مدى قدرة الأفراد والمؤسسات على إيجاد طرق لاستثمار التقنية بكفاءة، والقدرة على استيعاب التغيرات الحتمية، بالإضافة إلى القدرة على خلق الأطر الاجتماعية اللازمة لذلك التغير.

ومن المعروف أن عالم المعلومات عالم تتضافر فيه جهود الكثيرين من المساهمين، وكل منهم يسهم بخبرته وتجربته وموارده؛ ولإحداث تغيرات جوهرية في النظام فإن الأمر يتطلب تغيرات متبادلة في العلاقات الاقتصادية والاجتماعية والقانونية بين هؤلاء الشركاء.

ولما كانت المكتبات الرقمية تعتمد على البشر، فإنه لا يمكن بأي حال أن يتم تقديمها لهم بشكل أسرع من تبني هؤلاء ومؤسساتهم لهذه التقنيات، وينطبق هذا القول على كل من منشئي المكتبات الرقمية والمستفيدين منها والمهنيين الذين يدعمون أنشطتهم، مع مراعاة أن العلاقات بين هؤلاء قد تغيرت بالفعل؛ فمع المكتبات الرقمية يمكن للقراء أن يصلوا مباشرة إلى المعلومات دون زيارة لمبنى المكتبة وبدون أية اتصالات مع الوسطاء المهنيين، كما يمكن للمؤلفين أن يجروا المزيد من مسودات بحوثهم، ولكي يتمكن المهنيون من دعم كل تلك العلاقات الجديدة عليهم أن يكتسبوا مهارات جديدة، وأن يتلقوا تدريباً متقناً، وإذا كان من الممكن استثمار مثل هذه المهارات عن طريق الخبرة، فيمكن للآخرين أن يتعلموها. وتحاول كليات المكتبات الآن أن تكيف مناهجها مع هذه التغيرات، ولكن قد يستغرق ذلك وقتاً طويلاً قبل إحداث مثل تلك التغيرات المنشودة في النظام؛ حيث تأصلت التقاليد المكتسبة عبر مئات السنين.

وباستثناء بعض المجالات المتخصصة، فإن الحكمة تقول إن كلاً من المكتبات الرقمية والمجموعات التقليدية سوف تستمران في التعايش، وهذا يعني أن مكتبات المؤسسات سوف تظل محتفظة بمجموعات كبيرة من المواد التقليدية جنباً إلى جنب مع خدماتها الرقمية، كما أن الناشرين سوف يستمرون في المحافظة على أسواق كبيرة لمنتجاتهم التقليدية، ومع ذلك فإن الانخفاض المذهل والمستمر في تكلفة الحاسبات الآلية وما يصاحب ذلك من الزيادة المقابلة في قدرات هذه الحاسبات ستدفعان إلى البحث عن بديل كامل، والدليل على ذلك أن معالجات النصوص قد حلت محل الآلات الكاتبة في أقل من عشر سنوات، وكذلك الحال بالنسبة لبطاقات الفهارس التقليدية بالمكتبات، ففي

عام ١٩٨٠م لم يكن هناك سوى عدة مئات من المكتبات التي أعدت فهارسها الآلية المباشرة online catalogs، وفي غضون عشر سنوات لم تعد بطاقات الفهارس في الغالبية العظمى من المكتبات الأمريكية سوى تحف تاريخية، كذلك يمكن القول الآن إن المكتبات الرقمية يمكن أن تحل في بعض المجالات المتخصصة محل المكتبة التقليدية تماماً.

إن العديد من التطورات المثيرة في المكتبات الرقمية قد قدمتها مؤسسات جديدة، ويمكن لهذه المؤسسات الجديدة أن تبدأ من جديد، في حين يجب على المؤسسات الأقدم أن تعمل على المحافظة على الخدمات التي تقدمها أثناء تقديمها للخدمات الجديدة. ومن أهم التأثيرات المحتملة للمكتبات الرقمية هو التحويل الكامل للموارد المالية من حوزة الموردين التقليدين للمعلومات إلى متعهدي المعلومات الجدد وإلى صناعة الحاسبات، ومن الطبيعي أن المؤسسات الموجودة ستحاول جاهدة أن تحبط أي تغير يقلل من أهميتها، ولكن العلاقات الاقتصادية بين الأطراف قد تغيرت بالفعل. وسوف تقلص بعض المؤسسات من حجم تواجدها، بل قد تعلن إفلاسها. وإن التنبؤ بمثل هذه التغيرات قد يكون صعباً في ظل مظاهر التردد في تمويل المكتبات الرقمية وصناعة النشر الإلكتروني، وفي ظل غياب النظام القانوني الذي يحكم هذا التبني، ولكن في نهاية الأمر يمكن القول إن ضغط السوق سوف يخلق نظاماً جديداً، وفي وقت ما سوف تستقر السوق بشكل كاف لتفسح المجال ولتتضح صورة المبادئ القانونية اللازمة. وحتى ذلكم الحين فإن مظاهر التردد الاقتصادي والقانوني ستكون مقلقة، ومع ذلك فليس هناك من دليل على أنها ستكون عقبات جسورة في المستقبل.

وعلى المدى البعيد لا يبدو أن ثمة عقبات في طريق المكتبات الرقمية وصناعة النشر الإلكتروني؛ وبرغم كثرة التحديات الفنية والاقتصادية والاجتماعية والقانونية فإنه سيتم التغلب عليها بشكل مطرد. ومع أننا لسنا على يقين بطبيعة الشكل الذي ستكون عليه المكتبات الرقمية، فمن الواضح أنها وجدت لتستمر.

الفصل الثاني
الإنترنت ونسيج الشبكة العنكبوتية

الإنترنت :

الإنترنت، كما يوحي اسمها، مجموعة مترابطة من الشبكات التي تدار كل منها بشكل مستقل، كما أن كل شبكة من هذه الشبكات تدعم المواصفات الفنية القياسية اللازمة لعملية الترابط - وهي المواصفات التي تكون عائلة بروتوكولات التحكم بالإرسال / بروتوكولات الإنترنت (TCP/IP). بالإضافة إلى إحدى الطرق الموحدة للتعرف إلى الحاسبات الآلية - غير أنه في حالات كثيرة تكون تلك الشبكات المستقلة مختلفة تماماً. وتستخدم الأجزاء المختلفة من الإنترنت كل أنواع قنوات الاتصال تقريباً التي يمكنها نقل البيانات أو إرسالها، وتتراوح هذه القنوات بين القنوات السريعة والموثوق فيها إلى القنوات البطيئة والتي لا يعول عليها. كما أن هذه القنوات قد تكون قنوات خاصة أو عامة، ويتم دفع رسوم التعامل مع هذه القنوات بأساليب مختلفة. وتعرف الإنترنت في بعض الأحيان "بالطريق السريع للمعلومات Information Highway"، ومن أفضل ما يمكن مقارنتها به، ذلكم النظام الخاص بالنقل الدولي، الذي يشتمل على الكثير من النظم بدءاً من خطوط الطيران إلى خطوط السكك الحديدية الوعرة.

لقد نشأت الشبكات التي شكلت الإنترنت بأسلوبين : وكان أحد خطوط التطوير متمثلاً في الشبكات المحلية (Local Area Network (LAN التي أنشئت لربط مجموعة من الحاسبات الآلية والنهايات الطرفية الموجودة داخل إطار إحدى المصالح أو المؤسسات. وقد استمدت معظم المفاهيم

المرتبطة بهذا الأسلوب من جهود مركز بحوث بالو آلتو Palo Alto Research Center (PARC) التابع لشركة زيروكس.Xerox Corporation وقد كانت الجامعات بالولايات المتحدة الأمريكية هي الرائدة في توسعة مجال الشبكات المحلية الصغيرة إلى الشبكات الموسعة داخل المدن الجامعية Campus-Wide Network. أما المصدر الثاني في تطوير الشبكات فقد تمثل في الشبكات الإقليمية الموسعة .National Wide–Area Networks. وتعد شبكة آربانت ARPAnet أشهر الأمثلة على ذلك، وهي الشبكة التي استطاعت في منتصف الثمانينات من القرن العشرين أن تربط بين ما يقرب من مائة وخمسين منظمة من منظمات البحوث العلمية المتخصصة في مجال الحاسب الآلي. ومن الشبكات العملاقة كذلك شبكة بتنت Bitnet التي تكاد تنسى، والتي ورثت الإنترنت الحالية خدمة المستفيدين منها.

ولقد كانت عملية ربط الشبكات بعضها مع بعض صعبة في بدايتها، حيث كانت تستخدم طرقًا فنية متفاوتة في ذلك. لكن الموقف تغير في أواخر الثمانينات من القرن العشرين حيث تلاقت رؤى رجال الجامعات وغيرها من المجتمعات العلمية على استخدام بروتوكولات التحكم بالإرسال / بروتوكولات الإنترنت (TCP/IP)، والتي كانت بمثابة بروتوكولات شبكة آربانت، لتنشأ بذلك الإنترنت التي نعرفها اليوم. ويعد القرار الذي اتخذته المؤسسة الوطنية للعلوم (National Science Foundation(NSF في عام ١٩٨٦م الحدث الأساس في هذا الصدد، حيث قضى هذا القرار بإنشاء شبكة أساسية Backbone فائقة السرعة تخدم الولايات المتحدة الأمريكية وتدعم تطوير الشبكات الإقليمية المرتبطة بها. وقد قررت المؤسسة في سبيل إنشائها لتلك الشبكة الأساسية أن تتخذ من

الإنجازات الفنية التي حققتها شبكة آربانت معياراً قياسياً تعتمد عليه، وقد كان ذلك أساساً لشبكة الإنترنت. وفي هذه الأثناء كانت شبكات المدن الجامعية Campus Networks تستخدم تلك المواصفات نفسها، وتعد شبكة جامعة كارنيجي ميلون Carnegie Mellon فائقة السرعة التي اكتملت بنيتها في عام ١٩٨٦م هي أقرب الأمثلة على باكورة مثل تلك الشبكات، حيث استطاعت أن تعمل على تهيئة العديد من طرق نقل البيانات، ودعم جهود التنافس لإعداد بروتوكولات الشبكات التي قامت بها كل من شركة ديجيتل Digital، وآبل Apple، وآي بي إم IPM، غير أن القاسم المشترك كان استخدام بروتوكولات التحكم بالإرسال / وبروتوكولات الإنترنت (TCP/IP).

وربما لا يمكن للمؤرخين في المستقبل أن يتفقوا بشأن مجموعة العوامل المالية والتنظيمية والفنية التي أدت إلى قبول المواصفات القياسية الفنية الخاصة بشبكة آربانت؛ فهناك العديد من الشركات التي قدمت مساهمات كان لها أثر كبير في تطوير الإنترنت وفي توسعة مجالها، غير أن الريادة في هذا الصدد كانت لجهتين حكوميتين في الولايات المتحدة، هما: هيئة مشروعات البحوث المتطورة للدفاع DARPA، والمؤسسة الوطنية للعلوم (NSF).

إن تقنية الإنترنت هي موضوع لكتب كاملة، غير أن هناك كثيراً من التفاصيل التي تبدو قليلة الأهمية للمستفيدين، إلا أن فهم أساسيات هذه التقنية يعتبر مفيداً لكل من مصممي المكتبات الرقمية والمستفيدين منها. وتتضمن اللوحة رقم (٢-١) تعريفاً ببروتوكولات التحكم بالإرسال، وبروتوكولات الإنترنت (TCP/IP)، وبالعناوين الرقميةAddress Numeric المعروفة بعناوين

بـروتوكولات الإنتـرنت IP Addresses. وهـناك طـريقة أخـرى لـتحديد أحـد الحاسـبات علـى الإنتـرنت، وهـي إعطـاؤه اسـماً علـى الـنحو التالـي: Tulip.Mercury.Cmu.edu. وتعـرف الأسـماء بهـذا الـشكل بأسـماء الـنطاق Domain Names، كمـا يعـرف الـنظام الـذي يـربطها بعـناوين بـروتوكولات الإنتـرنت "بـنظام اسم النطاق Domain Name System (DNS)". وفـي العادة تقـوم الحاسـبات الآلـية التـي تـدعم بـروتوكولات الـتحكم بالإرسـال، وبـروتوكولات الإنتـرنت (TCP/IP) بتوفيـر مجموعة من المواصفات القياسية للتطبـيقات الأساسية، وهذه التطبـيقات تعـرف بمجمـوعة بـروتوكول الـتحكم بالإرسـال، وبـروتوكول الاتـصال المـشترك بـين الـشبكات TCP/IP suite. وتتضمن اللوحة رقم (٢-٢) بعضاً من أشهر تلك البروتوكولات المستخدمة.

اللوحة رقم (٢ - ١)

بروتوكول التحكم بالإرسال/ وبروتوكول الإنترنت (١) **TCP / IP**

يعد كل من بـروتوكول التحكم بالإرسال TCP، وبروتوكول الإنترنت IP البروتوكولان الأساسيان اللذان يكونان الإنترنت، ودائماً ما يظهر الاختصاران المعبران عن هذين البـروتوكولين متلازمين بدرجـة قد لا نتذكر بسهولة بأن هذين الاختصارين هما لبروتوكولين اثنين منفصلين بعضهما عن بعض.

ويعمل بـروتوكول الإنتـرنت (IP) علـى ربط أجزاء الشبكة المنفصلة التي تكوّن الإنتـرنت، فكل حاسـب آلـي على الإنترنت له عنوانه الفريد الذي يعرف بعنوان بروتوكول الإنترنت IP Address، وهذا العنوان يتكون من أربع شرائح رقمية (كمـا فـي المثـال التالـي : 90 .3 .132.151) وكل شريحة من هذه الشرائح

(١) يعرف أحياناً ببروتوكول الاتصال المشترك بين الشبكات (المترجمان) .

تتراوح الأرقام فيها بين صفر - ٢٥٥، ويتم اختزان هذه الشرائح داخل الحاسب كأربعة بتات Bites. وعندما تكتب العناوين جرت العادة على أن يفصل بين كل شريحة وأخرى بنقطة، كما في المثال السابق. ويعمل بروتوكول الإنترنت على تمكين أي حاسب على الشبكة من إرسال رسالة ما إلى أي حاسب آخر. وتتصل الأجزاء المختلفة من الإنترنت عن طريق حاسبات خاصة تسمى "موجهات المسارات Routers"، وتستخدم هذه الموجهات - كما يوحي اسمها - عنوان بروتوكول الإنترنت في توجيه كل رسالة في المرحلة الثانية من رحلتها إلى مقصدها المحدد.

وعادة ما ترسل الرسائل على الإنترنت كحزم packets قصيرة أشبه ببضع مئات من البايتات في طولها، وببساطة يقوم الموجه باستقبال حزمة البيانات من كل وحدة من وحدات الشبكة، ثم يعيد توجيهها إلى طريقها الذي تريد أن تسلكه. ومن الملاحظ أنه ليس لدى موجه مسارات بروتوكول الإنترنت IP Router أي طريقة يعرف من خلالها ما إذا كانت هذه الحزمة قد وصلت إلى مقصدها النهائي أم لا، ونادراً ما يهتم مستخدمو الشبكة بتلك الحزم الفردية أو بوحدات الشبكة. وإنما هم بحاجة إلى موزع على قدر كبير من الثقة يعمل على نقل الرسائل الكاملة من حاسب إلى آخر، وهذه هي وظيفة بروتوكول التحكم بالإرسال Transport Control Protocol (TCP)؛ حيث يقوم برنامج التطبيقات الموجود على الحاسب المرسل بتمرير رسالة ما إلى برمجية بروتوكول التحكم بالإرسال المحلي، ثم يقوم بروتوكول التحكم بالإرسال بأخذ الرسالة ثم تقسيمها إلى حزم وعنونة كل حزمة بعنوان بروتوكول الإنترنت للجهة التي تقصدها الرسالة مع إعطائها رقماً مسلسلاً، ثم إرسالها عبر الشبكة. وعند وصول الرسالة إلى الحاسب المستقبل، يتم

التعرف إلى كل حزمة عند استلامها، ثم إعادة تجميع الحزم في رسالة واحدة، ثم تسلم بدورها إلى برنامج التطبيق.

إن بروتوكول التحكم بالإرسال يجب أن يكون غير مرئي بالنسبة للمستفيدين من المكتبات الرقمية، غير أن استجابة الشبكة تتأثر بشكل كبير بالبروتوكول، وهذا غالباً ما يؤثر في الأداء النهائي الذي يلحظه المستفيد. فمن الملاحظ أن الحزم المرسلة لا تصل جميعها بنجاح، فالموجه المثقل بالمهام قد يتجاهل بعض هذه الحزم (أو يسقطها)، وإذا حدث ذلك فإن الحاسب المرسل لا يتلقى مطلقاً إشعاراً بالاستلام. وفي نهاية الأمر يمل الحاسب من طول الانتظار ويضطر إلى إعادة إرسال الحزمة مرة أخرى. وهذا ما يعرف بالوقت الضائع أو بفترة المهلة Time-Out والذي ينظر إليها المستفيد على أنه تأخير مزعج.

إن بروتوكول التحكم بالإرسال يضمن التوصيل الدقيق أو الخالي من الأخطاء error-free delivery للرسائل، ولكنه لا يضمن وصولاً عاجلاً لها، مع أن الوصول العاجل قد يكون في بعض التطبيقات أكثر أهمية من الدقة الكاملة. فلتفترض مثلاً أن أحد الحاسبات أرسل حديثاً مسموعاً، ثم قام الحاسب الآخر بإذاعته مباشرة بمجرد وصوله، فإذا لم تصل إحدى الحزم في الوقت المحدد، فإن المستمع كثيراً ما يفضل أن يتجاهل تلك الأجزاء الصغيرة جداً من مسار الصوت بدلاً من انتظار تلك الحزم المفقودة حتى يعاد إرسالها، والتي ستكون في غير سياقها وتصل في شكل مضحك أو غير صحيح، ولأن بروتوكول التحكم بالإرسال غير مناسب لمثل هذه التطبيقات، فإنه يتم استخدام بروتوكول بديل، ويعرف هذا البروتوكول بـ "يو دي بي UDP". وهو يعمل كذلك مع بروتوكول الإنترنت (IP)، وباستخدام بروتوكول "يو دي بي" هذا، يقوم

الحاسب المِرسل بإرسال سلسلة من الحزم على أمل أن تصل، ثم يقوم ببذل أقصى جهده في ذلك ولكنه لا يضمن وصول أي من هذه الحزم.

اللوحة رقم (٢– ٢)

مجموعة بروتوكولات التحكم بالإرسال وبروتوكولات الإنترنت

The TCP/IP Suite

مجموعة بروتوكولات التحكم بالإرسال وبروتوكولات الإنترنت هي مجموعة من برامج آلية (معتمدة على بروتوكولات التحكم بالإرسال، وبروتوكولات الإنترنت) تقدمها الحاسبات الأكثر حداثة، وهي تضم البرامج التالية :

- **محاكاة المحطة الطرفية Terminal Emulation :** وهو برنامج يعرف بالتيلنت Telnet يسمح لأحد الحاسبات الشخصية أن يحاكي محطة طرفية لحاسب من طراز قديم ليس له قوة تجهيز خاصة به، ودائماً ما يعتمد على حاسب بعيد في عملية معالجة البيانات. كما أنه عادة ما يستخدم برنامج التيلنت في إدارة النظام، لأن الحاسب القديم يوفر واجهة تعامل للمستفيد من أقل المستويات قبولاً.
- **نقل الملفات File Transfer :** يعرف البروتوكول الذي يقوم بعملية نقل الملفات من حاسب آلي إلى آخر عبر الإنترنت ببروتوكول نقل الملفات File Transfer Protocol (FTP)، وبما أن بروتوكول نقل الملفات قد صمم لاستخدام بروتوكولات التحكم بالإرسال، فإنه يعد فعالاً في عملية نقل الملفات الكبيرة عبر الإنترنت.
- **البريد الإلكتروني Electronic Mail :** يستخدم بريد الإنترنت بروتوكول نقل البريد المبسط Simple Mail Transport Protocol (SMTP)

وهو البروتوكول الذي يتولى مهمة تحويل البريد الإلكتروني من مجموعة خدمات محلية إلى خدمة عالمية فردية. وهي بدورها توفر الآلية الأساسية لعملية توصيل البريد، وحديثاً تم عمل سلسلة من البرامج الإضافية التكميلية Extensions التي تسمح للرسائل بنقل مجموعات من الرموز الكبيرة، ونقل البريد متعدد الوسائط، ودعم عملية إرفاق ملفات مع الرسائل البريدية.

اللوحة رقم (٢ – ٣)

أخبار الشبكة (نت نيوز) **Net News**

تعد لوحات نشرات أخبار الشبكة (والتي تعرف أيضاً بشبكة المستفيدين Usenet) مهمة، وهي نماذج واضحة لتوجه مجتمع الإنترنت إلى التوزيع الحر أو المفتوح للمعلومات. ويتم تنظيم آلاف من لوحات النشرات التي تسمى بالمجموعات الإخبارية Newsgroups في سلاسل هرمية، وتضم التجمعات عالية المستوى مجموعات إخبارية مهتمة بموضوعات مثل الحاسب الآلي .comp، والتسجيلات الصوتية، rec والبدائل المشهورة alt Motorions، وغيرها كثير(١). فعلى سبيل المثال تعد اللوحة الإخبارية المعنونة بـ rec.arts.theatve.musicals، إحدى لوحات النشرات لمناقشة القضايا الموسيقية.

ونظام أخبار الشبكة هو نظام لا مركزي إلى أبعد حد لدرجة أنه ليس هناك شخص واحد يمتلك قائمة شاملة بكل المجموعات الإخبارية. ومن ثم

---

(١) تضم المجموعات الإخبارية لشبكة المستفيدين Usenet Newsgroup تسلسلاً هرمياً للمجموعات الإخبارية على الإنترنت، يضم سبع مجموعات تتألف من comp., musc., news., rec., sci., soc., and talk (المترجمان عن حسب الله، ص ١١١).

فإن أي شخص يرغب في إرسال إحدى الرسائل إلى مجموعة معينة يرسلها إلى الحاسب المضيف المحلي للأخبار. ثم يقوم هذا الحاسب المضيف بدوره بتمريرها إلى جيرانه، الذين يقومون بدورهم بتمريرها إلى جيرانهم... وهكذا.

وتعد أخبار الشبكة Net News عكس المكتبة الرقمية بالضبط؛ حيث لا تخضع معلومات أخبار الشبكة لأي نوع من أنواع الإدارة، وليس هناك بالضرورة قيود على من يرسل المعلومات أو على طبيعة المعلومات التي يمكن للشخص أن يرسلها. ويقوم هذا النظام في أسوأ حالاته ببث المعلومات البذيئة وغير الأخلاقية، والصور الخليعة، أو على الأقل المعلومات غير الصحيحة، مع أن هناك العديد من المجموعات الإخبارية التي تعمل بشكل جاد. فعلى سبيل المثال يرتبط كثير من المستخدمين للغة بايثون للبرمجة Python Programming Language من جميع أنحاء العالم بمجموعة إخبارية تعرف بـ (Comp.Lang. Python)، يقومون من خلالها بتبادل المعلومات الفنية وطرح التساؤلات والاتصال بمطور هذه اللغة.

مجتمع الإنترنت :

لقد مهدت الإنترنت الطريق أمام فكرة المواصفات القياسية المفتوحة open standards، ففي عام ١٩٩٧م، حصل كل من فينتون سيرف Vinton Cerf وروبرت كهن Robert Kahn، على الميدالية الوطنية للتقنية نظير إسهاماتهم في مجال الإنترنت. وقد أثنت التصريحات على أعمالهم حول بروتوكولات التحكم بالإرسال وبروتوكولات الإنترنت وأشارت إلى أنهما "لم يكونا رائدين في مجال التقنية فحسب، بل في أسلوبهما الاقتصادي والفعال في نقل تلك

التقنية"، كما قيل كذلك بأنهما: "قد حافظا بمثابرة وصمود على أن تكون بروتوكولاتهم للربط الشبكي Internetworking Protocols متاحة مجاناً لأي شخص. وتستطرد التصريحات قائلة: "بأن بروتوكولات التحكم بالإرسال، وبروتوكولات الإنترنت (TCP/IP) صممت بتأن لتوزع بشكل مستقل ولدعم الشبكات عبر جميع خطوط الحاسبات، وجميع أساليب الإرسال".

وتستمر أعراف الإنترنت في التأكيد على التعاون في المجالات الفنية، وأن التطور الدائم للإنترنت يبقى رهن جهود المهندسين. وإن كان بعض الناس لا يزالون مترددين في تقبل المقولة بأن حكومة الولايات المتحدة الأمريكية مؤهلة لعمل أي شيء ذي شأن في هذا الاتجاه، مع أن نشأة الإنترنت كانت تقودها الجهات الحكومية. وغالباً ما كان يتم ذلك في مواجهة المقاومة القوية للشركات التي تجني الآن أرباحاً من وراء تلك النجاحات. وحديثاً جرت محاولات كثيرة لإعادة كتابة تاريخ الإنترنت من أجل دفع قيمة الفوائد المكتسبة، بل لقد ادعى أناس مسئوليتهم عن الإنجازات التي تحققت بفضل تضافر جهود الكثيرين. وهناك تباين واضح بين الترابط المنطقي للإنترنت - الذي قام بتنسيقه مسؤولون حكوميون ثاقبو النظر - وطوفان المواصفات القياسية غير المتوافقة في مجالات تركت ساحاتها للمنافسة التجارية، كما هو الحال في مجال الهواتف المحمولة.

والسمة الهامة المرتبطة بالإنترنت هي أن المهندسين وعلماء الحاسب الآلي الذين قاموا بتطويرها وتشغيلها هم أنفسهم أكثر المستفيدين من خدماتها. حيث يتواصلون عبر البريد الإلكتروني، متناسين بذلك البريد التقليدي الذي أصبح في نظرهم أشبه بالسلحفاة في سرعتها. كما أنهم يعتمدون على

حاسباتهم الشخصية عند تأليف بحوثهم بكتابتها وتحميلها عليها، وإذا كانت بحوثهم على شكل صفحات على الويب فإنهم يقومون بأنفسهم بإدخال تيجان تهيئة تلك النصوص Markup Tags بدلاً من استخدام أي برنامج للتنسيق Formatting program. كذلك يقضي كبار علماء الحاسبات أوقاتاً طويلة في إعداد عروض تقديمية عامة بدلاً من كتابة برامج آلية، مع أن البرمجة تعد مهارة أساسية يتوقع أن يكتسبها كل شخص.

النشر العلمي عبر الإنترنت :

يرجع تاريخ نشر الأعمال الأكاديمية الجادة على الإنترنت لعدة سنوات مضت، وتتضمن اللوحتان (٢-٤) و (٢-٥) وصفاً لنموذجين هامين، هما: سلسلة آر إف سي RFC، وأرشيفات الطباعة الإلكترونية لعلم الفيزياء Physics E-Print Archives في مختبر لوس ألاموس الوطني Los Alamos National Laboratory، وهما من النماذج غير المشهورة. وإذا كان الاختصار RFC يعني فيما مضى "طلب تعليق Request for Comment"، فإن سلسلة "آر إف سي RFC" تعد الآن السلسلة الفنية المميزة للإنترنت؛ حيث تتضمن مجموعة متنوعة من المعلومات الفنية والمواصفات القياسية الرسمية للإنترنت. أما خدمة لوس ألاموس فليست أرشيفاً بالمعنى المعروف للكلمة، وإنما تتمثل وظيفتها الأساسية في أنها تعمل كحاسبات خادمة للطبعات المبدئية من البحوث العلمية Preprint Server، بمعنى أنها موقع يمكن للباحثين أن ينشروا من خلاله أبحاثهم بمجرد الانتهاء منها، تجنباً للتأخير الذي تتصف به منافذ النشر التقليدية للدوريات العلمية.

وأياً كانت مضامين أسماء هاتين الخدمتين، فإنهما يحظيان بأهمية

خاصة بالنسبة لنشر البحوث العلمية، كما أنهما يكتسبان أهمية أخرى في تجسيدهما للسبل الجديدة التي أوجدتها المكتبات الرقمية لإنجاز المهام. وإذا كانت إحدى مقالات النشر العلمي تقول إن جودة ما يتم نشره لا يمكن أن تتحقق إلا من خلال عملية التحكيم ـ تلك العملية التي يتم من خلالها قراءة المقالة من قبل متخصص آخر غير مؤلفها قبل نشرها. فإن العملية التي أصبحت من خلالها مسودات الإنترنت بمثابة "طلبات تعليق RFCs"، تعد شكلاً قويًا من أشكال التحكيم، غير أن هذه العملية غالباً ما تتم بعد نشر مسودة المقالة بشكل رسمي. ومع أن خدمة لوس ألاموس Los Alamos لا تتضمن عملية تحكيم، فقد ثبت أنهما من طرق الاتصال العلمي الفعالة إلى حد كبير.

وهاتان الخدمتان فاعلتان من الناحية الاقتصادية؛ حيث تفتحان أبوابهما على مصراعيهما للمستفيد، ومع أن تشغيلهما يتم بشكل مهني، وبميزانيات ضخمة، فليس ثمة أية رسوم تفرض على المؤلفين الذين ينتجون المعلومات أو على القراء الذين يفيدون من تلك المعلومات.

ومع أن تلك الخدمتين المشار إليهما في الفقرات السابقة قد عرفتا قبل نشأة الويب، فإن الويب سجلت نجاحاً باهراً أنسى الكثيرين أن ثمة طرقاً أخرى فعالة لبث المعلومات عبر الإنترنت. وبالرغم من أن أرشيفات لوس ألاموس، وسلسلة طلبات التعليق RFC، تستخدمان الآن أساليب الويب، فإنهما اعتمدتا في نشأتهما على تقنيات البريد الإلكتروني ونقل الملف.

اللوحة رقم (٢ ـ ٤)

لجنة العمل الخاصة بهندسة الإنترنت وسلسلة طلب التعليق

The Internet Engineering Task Force and the RFC Series

لجنة العمل الخاصة بهندسة الإنترنت هي الجهاز الذي يضطلع بمهام التنسيق بين الجوانب الفنية للإنترنت. وتبدو أساليب العمل فيها فريدة، وقد ثبت حتى الآن أنها جيدة بشكل غير معتاد لجمعها أعداداً كبيرة من المهتمين الذي ينتسب الكثير منهم لشركات متنافسة، للعمل سوياً. ولعل أول مزايا هذه اللجنة الخاصة بهندسة الإنترنت أن أبوابها مفتوحة للجميع، فأي شخص يمكنه حضور اجتماعاتها، ويشترك مع مجموعات العمل، بل ويدلي بصوته.

وترفع لجنة هندسة الإنترنت مقولة "الإجماع التقريبي وميثاق العمل Rough Consensus and Working Code" شعاراً أساسياً لعملها، بحيث يشجع أي شخص لديه اقتراح لبروتوكول جديد، أو لديه فكرة لأي تطور فني آخر، فإنه يُطلب منه أن يقدم تقريراً فنياً Technical Paper (يسمى مسودة إنترنت An Internet Draft) وتصوراً عملياً لتنفيذ الفكرة Reference Implementation ، على أن يكون هذا التصور في شكل برمجية متاحة مجاناً، بحيث يتم مناقشة هذه المسودة في لقاءات مجموعات العمل، وإذا ما كان هناك إجماع على وجاهة الفكرة وصلاحيتها بعد المناقشة، يتم وضعها ضمن مسار مواصفات أو معايير طلب التعليق RFC Standards، ولا يمكن للمسودة أن تصبح مواصفة قياسية رسمية إلا بعد توفير طريقة تنفيذها للجميع (عادة ما تكون في شكل برامج على الحاسب الآلي).

ومع أن لجنة هندسة الإنترنت قد بدأت نشاطها في الولايات المتحدة، فإنها اتخذت طابعاً دولياً الآن؛ حيث يعقد لقاء سنوي لها خارج حدود الولايات المتحدة، ويحضره المشاركون بمن فيهم رؤساء جماعات العمل من

جميع أنحاء العالم. ومع أن هذه اللجنة كانت تمول في الأساس من قبل برنامج المنح الذي تقدمه حكومة الولايات المتحدة، فقد تحقق لها الآن اكتفاءً ذاتياً إذ تتكفل رسوم حضور اجتماعاتها بتغطية التكاليف.

إن أوجه العمل في لجنة هندسة الإنترنت مفتوحة لجميع من يرغب في المساهمة، على عكس بعض جهات المواصفات القياسية الأخرى التي يصعب الحصول على مسوداتها. فضلاً عن أن مواصفاتها عادة ما تكون معقدة وباهظة الثمن، فإن مسودات الإنترنت وطلبات التعليق تتاح مباشرة On-Line، ولأن التركيز ينصب على برمجيات الأعمال، فإن أول أسلوب من الأسلوبين الفنيين المتنافسين المقدمين مع البرمجية التي تعمل بالفعل، يلقى فرصة كبيرة للقبول. ونتيجة لذلك فإن المواصفات القياسية الأساسية للإنترنت بسيطة بدرجة لافتة للنظر.

وقد اجتذبت الاجتماعات للجنة هندسة الإنترنت أكثر من ألفي شخص، وقد خيمت روح المودة أو الألفة على أجواء اللقاءات وخاصة أن هؤلاء توزعوا على جماعات عمل يناقش كل منها قضايا محددة، وقد كان معظم هؤلاء مهندسين متمرسين ومتخصصين في مجال الحاسب الآلي، أما المديرون فقد مكثوا في منازلهم، حيث يرون أن اللقاءات الرسمية عادة ما تكون قصيرة وغير موثقة، أما اللقاءات غير الرسمية فهي طويلة ومجهدة. وقد أثمرت بعض الجلسات المسائية المتأخرة من جلسات اللجنة عن العديد من المواصفات القياسية الهامة بفضل تضافر جهود الكثيرين من العاملين بالمنظمات المتنافسة.

ويخشى على الإنترنت من خطر الوهن أو التفكك الفني بسبب نموها

المطرد، وتعد لجنة هندسة الإنترنت السبب الأساس في ذلك لأنها تبدي قدراً كبيراً من المرونة، فإذا تحكمت شركة وحيدة في الإنترنت، فإن تقنيتها سوف تتمشى مع مدى كفاءة الهيئة العليا لمهندسي تقنية الإنترنت، مع أن أفضل مهندسي العالم يعملون سوياً من أجل تدارك المشكلات المتوقعة.

ومسودات الإنترنت سلسلة جديرة بالملاحظة إذا نظر إليها كأحد المطبوعات الفنية. وبرغم التقادم السريع لمعظم المعلومات في مجال العلوم والهندسة، يتم وضع الدوريات على رفوف المكتبات إلى الأبد، وعلى عكس ذلك مسودات الإنترنت؛ حيث تصدر كل مسودة من هذه المسودات بعبارة ثابتة تقول: "إن مسودات الإنترنت ما هي إلا وثيقة أولية أو تمهيدية Draft Document تستمر صلاحيتها لمدة أقصاها ستة أشهر، كما يمكن لها أن تُحدّث أو تستبدل في أي وقت أو أن تستوعب في وثيقة أخرى، وإن من غير المناسب استخدام مسودة الإنترنت كمادة مرجعية يتم الاستشهاد بها، وحسبها في ذلك أنها قيد الإعداد ".

وتعمل لجنة هندسة الإنترنت على نشر كل مسودة إنترنت تقدم إليها مباشرة عبر الإنترنت On Line ومن ثم تبليغ المهتمين بها من خلال قوائم البريد الإلكتروني، وتبدأ عملية مراجعة هذه المسودات، ويقوم المهتمون بإرسال تعليقاتهم على هذه المسودات عبر قائمة البريد الإلكتروني، والتي تتراوح بين التعليقات التفصيلية إلى الانتقادات اللاذعة. وبعد ذلك تلتقي مجموعات العمل سوياً لمناقشة هذه المسودات في ضوء ما أثير حولها من تعليقات، وأخيراً يخضع الشكل المقترح للمراجعة في المرحلة النهائية من قبل الخبراء في هذا المجال.

وتعد طلبات التعليق RFC المطبوعات الرسمية للجنة هندسة الإنترنت، كما تشكل هذه المطبوعات التي تبلغ بضعة آلاف، والتي تعود إلى عام ١٩٦٩م، جوهر عملية التوثيق للإنترنت. ومن أشهر طلبات التعليق تلك التي تكوّن مسار المواصفات القياسية Standards Track، وهي تضم المواصفات الرسمية لكل إصدارة من بروتوكولات الإنترنت وبريد الإنترنت ومكونات نسيج الشبكة العنكبوتية وغيرها كثير، وهناك نوع آخر من طلبات التعليق وهو طلبات التعليق الإعلامية Informational RFCs التي تتولى مهمة نشر المعلومات الفنية المتصلة بالإنترنت.

ومما تجدر الإشارة إليه أن مناقشات قضايا النشر العلمي نادراً ما تشير إلى سلسلة طلبات التعليق، ولذلك فإن من الصعب أن نجد مجموعة أخرى من المطبوعات الهندسية أو العلمية تراجع بشكل مكثف قبل نشرها، أو تقرأ على نطاق واسع من قبل الخبراء في هذا المجال. ولم تنشر طلبات التعليق هذه مطلقاً في شكل مطبوع، وإنما تتاح عبر الإنترنت من خلال بروتوكول نقل الملفات، وهي متاحة الآن على الويب (ولكن لا تزال في صيغة الملفات النصية الأساسية Text-only format) [١].

اللوحة رقم (٢ - ٥)

أرشيفات لوس ألاموس للطباعة الإلكترونية

The Los Alamos E-Print Archives

تقدم أرشيفات الطباعة الإلكترونية في مجال الفيزياء نموذجاً واضحاً لاستثمار العلماء مزايا تقنية الإنترنت في تكوين شكل جديد من أشكال

(١) هي ملفات آسكي التي لا تخضع لأي نمط من التنسيق، ولا تتضمن أية رسومات أو مواد توضيحية إضافية (المترجمان).

الاتصال العلمي من خلال توسعة مجال تداول الطبعات المبدئية من البحوث العلمية. وقد أنشأ باول جنسبارك Paul Ginsparg بمختبر لوس ألاموس الوطني أول أرشيف في عام ١٩٩١م، بهدف تلبية احتياجات مجموعة من علماء فيزياء الجسيمات High-Level Physicists، ومؤخراً تم إنشاء الأرشيفات الأخرى لخدمة الفروع الأخرى لعلم الفيزياء، والرياضيات والمجالات المرتبطة بها، وقد قال باول جنسبارك في إحدى جلسات اليونسكو في باريس عام ١٩٩٦م: "إن هذه الأرشيفات تخدم الآن أكثر من ٣٥٠٠٠ مستفيد من أكثر من سبعين دولة من دول العالم، وتقوم يومياً بمعالجة أكثر من ٧٠٠٠٠ عملية إلكترونية Electronic Transactions، كما أنها قامت بالفعل في بعض مجالات الفيزياء باستئصال المجلات البحثية التقليدية كمنافذ لنقل المعلومات البحثية الجارية أو التاريخية من جذورها".

إن الوظيفة الأساسية لأرشيفات الطباعة الإلكترونية هي عرض نتائج البحوث العلمية، وغالباً ما يكون ذلك في شكل إصدارة تمهيدية لأحد البحوث الذي سينشر بعد ذلك في إحدى المجلات التقليدية. وعادة ما يتم إعداد المقالات للأرشيفات بالشكل المعتاد. ومع أن العديد من الفيزيائيين يستخدمون صيغة تي إكس T Ex [(١)] في كتابة هذه المقالات، فإن لغة بوست سكريبت Postscript، ولغة تهيئة النص الفائق HTML تستخدمان أيضاً في هذا الصدد، أما بالنسبة للرسومات والبيانات التي تشتمل عليها تلك المقالات فترد أحياناً مصاحبة للنص، وأحياناً أخرى ترد في ملفات مستقلة.

ويمكن أن يقدم البحث إلى الأرشيف عبر البريد الإلكتروني أو عن طريق

---

(١) منسق نصوص الرياضيات وضعه دونالد كنوث Donald/Cnuth حتى يمكن كتابة المعادلات والصيغ الرياضية في شكل رياضي (المترجمان) .

عملية نقل الملفات باستخدام بروتوكول نقل الملفات FTP أو من خلال الويب. ويتوقع من المؤلف أن يقدم ملخصاً قصيراً أو مجموعة مقننة من ما وراء البيانات اللازمة لعملية التكشيف Indexing Metadata، ثم تتم عملية المعالجة برمتها تلقائياً، وتقدم الأرشيفات للمشتركين خدمات البحث المعتمدة على البريد الإلكتروني، ونظم البحث المعتمدة على الويب، وخدمات الإعلام بالبريد الإلكتروني. كما تضم خيارات البحث عدة بدائل، منها: البحث في أرشيف واحد، والبحث في عدة أرشيفات، والبحث بالمؤلف والعنوان، وبحث النصوص الكاملة للمستخلصات.

ولا شك أن تقنية الأرشيفات هي تقنية واضحة المعالم؛ حيث تستخدم الصيغ القياسية والبروتوكولات ووسائل المشابكة التي يعرفها الباحثون ويفهمون مقاصدها. كما صممت واجهات تعامل المستفيدين بشكل يكفل تقليص الجهد اللازم لصيانة الأرشيف. ويتوقع المؤلفون والقراء الحصول على المساعدة من خلال تنصيب البرامج المناسبة على حاسباتهم الشخصية، فضلاً عن اتباعهم للإجراءات المطلوبة في هذا الصدد.

إن هذه الأرشيفات هي نظم إتاحة أو وصول مفتوحة Open-Access System، يتم تمويلها من خلال منحة سنوية تقدمها المؤسسة الوطنية للعلوم National Science Foundation، ووزارة الطاقة Department of Energy. وتتكفل هذه النظم بحماية حقوق المؤلفين، وتجدر الإشارة في هذا السياق إلى أن الإشارة الوحيدة تقريباً الخاصة بحق المؤلف توجد فقط في الصفحة الخاصة بإثبات الملكية الفكرية أو الصفحة التي تتضمن الشكر والتقدير.

وقد كتب جنسبارك قائلاً : "إن كثيراً من الدروس المستفادة من هذه

النظم يجب أن تنقل إلى مجالات أخرى من مجالات النشر العلمي Scholarly Publicationأعني بذلك تلك المجالات التي لا يكتب فيها المؤلفون بهدف الحصول على مكافآت مالية مباشرة نظير كتاباتهم، وإنما تلك المجالات التي يكتبون فيها بهدف تبادل المعلومات، من أجل تقدم المعرفة، بالإضافة إلى الثمرات التي تعود على وظائفهم وعلى مكاناتهم الوظيفية.

نسيج الشبكة العالمية The World Wide Web :

نسيج الشبكة العنكبوتية (المعروف بالويب) هو أحد النجاحات العظيمة في تاريخ الحاسبات الآلية، وقد نمت هذه الشبكة مع تطور معالجات النصوص والجداول الإلكترونية spreadsheets كأحد التطبيقات المميزة في مجال استخدامات الحاسبات الآلية. ولقد كانت الويب، وما ارتبط بها من تقنيات، نقطة تحول في رحلة النمو المتسارع للمكتبات الرقمية، ويسعى هذا الجزء من الكتاب إلى إلقاء نظرة سريعة أساسية على الويب وما يرتبط بها تقنيات، في حين تتكفل أجزاء أخرى من هذا الكتاب باستعراض التفاصيل عن جوانب أخرى محددة منها.

والويب هي مجموعة مترابطة من المعلومات المتوافرة في العديد من الحاسبات الإلكترونية المرتبطة بالإنترنت حول العالم، وتعرف هذه الحاسبات بخوادم [١] الويب Web Servers. ويقوم الأفراد على صيانة بعض خوادم الويب وما تحتويه من معلومات، كما يقوم بصيانة بعضها الآخر مجموعات صغيرة كالأقسام الأكاديمية بالجامعات أو مراكز البحوث في حين أن

(١) لعل الكلمة العربية ندل جمع نادل هي الأكثر دلالة في هذا السياق، غير أن كلمة "خادم" يشيع استخدامها بشكل أكبر (المترجمان) .

بعضها الآخر يُصان بواسطة شركات ضخمة كشركات خدمات المعلومات. وتجدر الإشارة إلى أن بعضاً من الموقع على الإنترنت صُممت كمكتبات رقمية، في حين يدار بعضها الآخر (القليل من هذه المواقع ممتاز) من قبل أشخاص لا يعتبرون أنفسهم مكتبيين أو ناشرين. كما أن بعض خوادم الويب تتوافر لـه مجموعات كبيرة من المعلومات القيمة ذات الجودة العالية، هذا في الوقت الذي تستخدم فيه بعض الخوادم الأخرى لفترات قصيرة أو لأغراض خاصة، وهي تدار على أية حال بشكل غير رسمي، أو تستخدم لأغراض تخرج عن إطار خدمات المكتبات كما هو الحال عند استخدامها لأغراض التسويق.

لقد قام تيم بيرنرز- لي Tim Berners -Lee وزملاؤه في المركز الأوربي للبحوث النووية (CERN) Center Europeans Pour La Recherche Nuclear في سويسرا بتطوير تقنية الويب في حدود عام ١٩٩٠م. وقد انتشرت بفضل برنامج الفسيفساء أو الموزايك Mosaic الذي يعد إحدى واجهات تعامل المستفيدين (١) User Interface، والذي قام بتطويره مارك أندرسن Marc Andresen وآخرون في جامعة إلينويز في أربانا - شامبين University of Iilinois at Urbana -Champaign. وكان هذا البرنامج - الموزايك - قد صدر وأطلق في أول الأمر في عام ١٩٩٣م، ثم تبعه في غضون سنوات قليلة إطلاق إصدارات تجارية أخرى منه، من أشهرها: متصفح نتسكيب Netscape Navigator، ومستكشف الإنترنت Microsoft Internet

(١) كثيرًا ما يصطلح على تسميتها بواجهات المستفيدين، أو واجهات تعامل المستفيدين (المترجمان).

Explorer. وتعرف واجهات المستفيدين هذه بمتصفحات الويب Web Browsers أو المتصفحات Browsers على سبيل التبسيط [١].

وبإيجاز شديد يمكن تلخيص السبب الرئيسي الذي كان وراء نجاح الويب في القول بأنها أتاحت طريقة ملائمة لبث المعلومات عبر الإنترنت، حيث يمكن للأفراد نشر المعلومات، وإتاحة سبل الوصول إليها أو التعامل معها بأنفسهم، من غير برامج تدريبية أو مساعدة من قبل الآخرين. وبمعرفة متواضعة باستخدام الحاسبات الآلية يمكن إنشاء مواقع على الويب،كما أنه وبهذه المعرفة المتواضعة يمكن استخدام أي من متصفحات الويب.

إن الويب كان ظاهرة أساسية، ليس فقط كتقنية مبسطة، وإنما بسبب الطريقة التي طرحت بها للعامة؛ حيث عملت على إزالة كافة الحواجز التي تعوق استخدامها. ويعد ما حدث في جامعة كارنيجي ميلون Carnegie Mellon University خير شاهد على ذلك؛ حيث قام الموظفون باستنساخ برمجية الويب الموجودة على الإنترنت وتحميلها على حاسباتهم الخاصة، ثم قاموا بعد ذلك بتحميل المعلومات التي تدخل في إطار اهتماماتهم وإتاحتها للآخرين، وقد كانت حاسباتهم متصلة بشبكة المدينة الجامعية The Campus Network، ومن ثم بالإنترنت. وخلال ستة أشهر من الإطلاق الأول لبرنامج الموزايك، تمكن ثلاثة أشخاص من إنشاء ثلاث مجموعات ضخمة من المعلومات العلمية في مجالات الإحصاء، واللغة الإنجليزية وآدابها، وعلوم

(١) هناك تسميات عربية أخرى منها برامج استعراض الويب أو المستعرضات (المترجمان).

البيئة. وبحكم الاتساع العالمي للإنترنت، تمكن كثيرون من الوصول إلى هذه المعلومات أو التعامل معها مباشرة. وفي نهاية المطاف رأت الجامعة أن تتبنى الويب بشكل رسمي، من خلال تصميم دقيق لصفحة رئيسة Home Page تتضمن معلومات عن الجامعة، ولم يتم ذلك إلا بعد أن عرف الأفراد طريقهم إلى ذلك.

ولعل السبب الأول الذي جعل الكثيرين قادرين على خوض التجارب مع الويب هو أن برمجية الويب هذه كانت متاحة دائماً مجاناً عبر الإنترنت، حيث طرح كل من المركز الأوربي للبحوث النووية CERN، وجامعة إلينويز مبدأ التوزيع المجاني لبرمجياتهما لكل من خوادم الويب وواجهات المستفيدين. ويعد برنامج آباشي Apache - الإصدارة المجانية من خادم الويب لجامعة إلينويز- من أشهر خوادم الويب استخدامًا اليوم، كما يتيح التوزيع المجاني للبرمجية على الإنترنت من ناحية أخرى حافزاً قوياً لتقبل التقنية الحديثة؛ فالغالبية العظمى من الفنيين يستمتعون بعملية التجريب في الوقت الذي يمقتون فيه البيروقراطية، بل يحرمون تلك المتعة والسعادة عندما يتوجب عليهم إصدار أمر شراء لتجاربهم، أو عندما يطلب من المدير أن يوقع على ترخيص البرمجية Software License.

وهناك سبب آخر كان وراء النجاح السريع الذي حققته الويب، وهو أن هذه التقنية عملت على توفير بوابات مرور Gateways للمعلومات التي لم تنتج أساساً لأغراض الإتاحة على الويب، صحيح أن المتصفحات قد صممت حول بروتوكول الويب المعروف ببروتوكول نقل النصوص الفائقة (http) Hypertext Transfer Protocol، ولكن هذه المتصفحات تقوم كذلك بدعم

بروتوكولات أخرى من بروتوكولات الإنترنت، مثل بروتوكول نقل الملفات (File Transfer Protocol (FTP، وأخبار الشبكة Net News، والبريد الإلكتروني، كما أن الدعم المقدم لكل من بروتوكولي جوفر Gopher، ووايز Wais، اللذين يكادان يختفيان الآن، قد سمح لمجموعات مبكرة من المعلومات أن تتعايش مع مواقع الويب الأولى. كما أن إحدى الآليات الأخرى، المعروفة بواجهة بوابة العبور المشتركة (Common Gateway Interface (CGI، قد سمحت للمتصفحات بسد الهوة التي كانت قائمة بين الويب وأي نظام آخر لاختزان المعلومات على الخط المباشر On- Line Information. وبهذه الطريقة فإن كميات كبيرة من المعلومات كانت متاحة بمجرد وجود برنامج الموزايك.

ومنذ الإطلاق الأول لبرنامج الموزايك، وُجدت المتصفحات الرائدة لمعظم نظم التشغيل الشائعة ـ والتي تتمثل في الإصدارات المختلفة من نظم النوافذ Windows، وماكنتوش Macintosh، و يونكس Unix ـ كما تُقدم هذه المتصفحات الآن لجميع الحاسبات الآلية القياسية. كما أن مدير موقع الويب يمكن أن يكون متيقناً من أن المستفيدين حول العالم سوف يشاهدون المعلومات المقدمة في الموقع بالشكل والهيئة ذاتهما تقريباً، مهما تكن أنواع الحاسبات الآلية المستخدمة من قبل هؤلاء المستفيدين.

تقنية الويب The Technology of the Web :

تعتمد الويب فنياً على أربعة أساليب مبسطة هي: لغة تهيئة النصوص الفائقة HTML، وبروتوكول نقل الملفات الفائقة HTTP، وأشكال بيانات مايم MIME، والمحددات الموحدة لمواقع المصادر URLs، ولكل واحد من هذه المفاهيم أهميته الخاصة، التي تتجاوز مجال الويب إلى المجال العام لعملية

التشغيل المتداخل Interoperability للمكتبات الرقمية.

لغة تهيئة النصوص الفائقة Hypertext Mark Up Language :

لغة تهيئة النصوص الفائقة هي لغة تهدف إلى وصف بنية الوثائق النصية ومظهرها، وتصور اللوحة رقم (٢-٦) أحد الملفات المبسطة المهيأة وفقاً للغة تهيئة النصوص الفائقة، وكيف يمكن لأي متصفح نموذجي أن يعرضها ويترجمها.

وكما يوضح المثال فإن الملف بلغة تهيئة النص الفائق يشتمل على كل من النص الذي سيتم ترجمته وتشفيره و التيجان Tags التي تصف شكل وبنية النص. وتيجان لغة تهيئة النصوص الفائقة تُعرف دائماً من خلال هذين المحددين "< and >". ويلاحظ أن معظم هذه التيجان عادة ما ترد في زوجين، مع تضمين الشرطة المائلة لتبين الزوج الثاني من التيجان، ومثال على ذلك "<title> and </title>" الذي يحتوي على نص يترجم ويعرض كعنوان، كما أن بعضاً من تيجان لغة تهيئة النصوص الفائقة تحدد الشكل الذي يجب أن يظهر فيه النص، كأن يظهر هذا النص بخط مائل، فيكتب ذلك على النحو التالي:

"<i> .... </i>"، كما أن إضافة التاج "
"ـ يشير إلى نهاية الفقرة على أن يبدأ النص من فقرة جديدة، وهناك تيجان أخرى تحدد بنية النص، مثل: <p>.....</p>، المحددين للفقرة، والتاجين <h1>.....</h1> ـ اللذين يشيران إلى أن النص المحصور بينهما يجب أن يرد كعنوان رئيسي من بداية الفقرة.

وتجدر الإشارة إلى أن التيجان البنائية لا تحدد الشكل، بل يترك ذلك للمتصفح. فعلى سبيل المثال تقوم كثير من المتصفحات بإدراج سطر خالٍ أو أبيض Blank Line في بداية الفقرة لتوحي بأنها فقرة جديدة، وهو على أية حال أسلوب شكلي يقرره المتصفح.

ويوضح المثال الذي تتضمنه اللوحة رقم (٢-٦) ميزتين أخريين خاصتين بلغة تهيئة النص الفائق، وهما من المزايا التي كان لهما دور حيوي في نجاح الويب، تتمثل الميزة الأولى في إمكانية إدراج صورة ملونة في صفحات الويب، فالتاج المكتوب على النحو: <img src="logo.gif"> هو توجيه إلى إدراج صورة مختزنة في ملف مستقل، مع ملاحظة أن الاختصار img يشير إلى الصورة Image، في حين يشير المختصر src إلى مصدر Source الصورة أو المكان الذي توجد فيه. كما تشير سلسلة الرموز التي تلي ذلك إلى اسم الملف الذي تختزن به الصورة. وبكتابة هذا الأمر المبسط، يقوم برنامج الموزايك بإحضار الصور الملونة وإتاحتها على صفحات الإنترنت. ومما تجدر الإشارة إليه في هذا الصدد أن تطبيقات الإنترنت قبل ظهور الويب كانت مملة ورتيبة، وكانت التطبيقات الشائعة تستخدم النصوص غير المنسقة أو غير موحدة الصيغ Unformatted والخالية من الصور. وقد كانت الويب بمثابة أول النظم المستخدمة على نطاق واسع للجمع بين النصوص المنسقة Formatted Text، والصور الملونة. وبذلك عادت الحياة إلى الإنترنت. أما الميزة الثانية والأكثر أهمية فهي استخدام الروابط الفائقة Hyperlinks؛ فمتصفحات الويب لا تعمل بمفردها، وإنما يمكن أن ترتبط أو تتصل مع صفحات أخرى في أي مكان على الإنترنت؛ فالمثال الموجود باللوحة (٢-٦)

يشتمل على رابط فائق واحد، وهو الذي يعبر عنه بالتاج التالي :

< a href=“http://www.dlib.org/dlib.html”>

وهذا التاج يتبعه نص مطول تحدد نهايته بالمحدد ، وعندما يعرض المتصفح هذا التاج، فإن هذا النص يبدو مظللاً؛ حيث يظهر في العادة بلون أزرق تحته خط. وهو تقليد سهل؛ بحيث يمكن لأي مستفيد أن يصل إلى الموقع بمجرد الضغط على هذا النص المكتوب باللون الأزرق الذي يوضع تحته خط. وهذه الدعوة إلى استحضار هذا الرابط عملية سهلة سواء بالنسبة للمستفيد، أو لمنشئ صفحة الويب، ففي هذا المثال يلاحظ أن الرابط لصفحة بلغة تهيئة النص الفائق موجودة على حاسب آخر، وهي الصفحة الرئيسية لمجلة المكتبات الرقمية D-Lib Magazine.

إن المفاهيم الأساسية للغة تهيئة النص الفائق يمكن تقديمها في جلسة تعليمية قصيرة، ويمكن استيعابها بسرعة، ومن خلال عملية تتويج مبسطة يمكن عمل وثيقة جذابة؛ ففي بداية الويب كان كل مستفيد يعلم نفسه بنفسه. يضاف إلى ذلك أن بناء الجملة في لغة تهيئة النص الفائق يتسم بالسهولة، في حين يتصف بالصعوبة في اللغات الحاسوبية الأخرى؛ لدرجة أن تجاهل فاصلة منقوطة مثلاً في إحدى تلك اللغات يمكن أن يسفر عن فشل البرنامج أو على الأقل إعطاء نتائج خاطئة. أما مع لغة تهيئة النص الفائق، ولو لم تكن عملية التهيئة قد تمت بشكل دقيق، فإن معظم المتصفحات سوف تتقبلها في الغالب، ولقد كانت هذه البساطة هي بمثابة القيمة المضافة، لأن برامج الحاسبات التي تستخدم في تفسير لغة تهيئة النص

الفائق وفي عرض صفحات الويب كانت سهلة في كتابتها.

اللوحة رقم (٢ - ٦)

مثال من صفحة مكتوبة بلغة تهيئة النص الفائق

Here is a simple text file in HTML format as it would be stored in a computer. It shows the use of tags to define the structure and format of a document.

```
<html>
<head>
<title>D-Lib</title>
</head>
<body>
<h1>D-LibMagazine</h1>
<img src = "logo.gif">
<p>Since the first issue appeared in July 1995,
<ahref="http://www.dlib.org/dlib.html">Dilbmagazine </a> has
appeared monthly as a compendium of research, news, and progress in
digital libraries.</p>

<p><i>WilliamY.Arms
<br>January 1, 1999</i></p>
</body>
</html>
```

When displayed by a browser, this document might be rendered as follows. The exact format depends upon the specific browser, computer,

and choice of options.



المحددات الموحدة لمواقع المصادر (Uniform Resource (URLs Locators) :

المحدد الموحد لموقع المصدر هو المكون الأساسي الثاني للويب، ومع أن هذه المحددات تبدو مزعجة، فإنها تتمتع بالمرونة؛ حيث تقدم آلية مبسطة لعمل العناوين addressing التي تسمح بدورها للويب بعمل روابط للمعلومات المحملة على الحاسبات في جميع أنحاء العالم. وهناك محدد موحد لموقع المصدر مبسط بالمثال الموجود في اللوحة رقم (٢-٦)، وهو Http=//www.dlib.org/dlib.html؛ حيث http هو اسم البروتوكول، و www.dlib.org هو اسم النطاق للحاسب، وdlib.html هو اسم الملف الموجود على ذلك الحاسب. ويمكن تفسير هذا المحدد الموحد من قبل المتصفح على النحو التالي : "باستخدام بروتوكول نقل الملفات الفائقة Http، صل الحاسب الآلي بعنوان الإنترنت الموسوم www.dlib.org، ثم احصل على الملف"dlib.htm ".

بـروتوكول نقـل الملفـات الفائقـة HTTP Hyper Text Transfer Protocol :

البـروتوكول فـي عـالم التحسيب يعنـي مجموعة القواعد التـي تستخدم فـي إرسـال الرسـائل بـين نظـم الحاسـبات. ويـشتمل البـروتوكول النموذجـي علـى وصـف للصيغ المستخدمة، والرسائل المختلفة، والتتابع المنطقي الذي يجب أن ترسل به هذه الرسائل، والاستجابات الملائمة، وحالات الأخطاء التي يمكن أن تحدث ... إلى آخر ذلك. كما أن بروتوكول نقل الملفات الفائقة هو البروتوكول المستخدم في إرسال الرسائل بين متصفحات الويب وخوادمها.

ويعد الأمر "توصل إلى أو احصل على Get" هو نمط الرسالة الأساسية فـي بـروتوكول نقل الملفـات الفائقـة، فبالـضغط علـى الـرابط الـذي يـشير إلـى المحـدد الموحد الموسـوم " Http =// www.dlib.org/dlib.html "، يعنـي تنفيذ الأمـر "توصـل إلـى" عـن طـريق بـروتوكول نقـل الملفـات الفائقـة http get command.، وهذا الأمر يفسر بشكل رسمي على النحو التالي: "أنشئ اتصال بـين المتـصفح وخـادم الويب الذي يحمل اسم النطاق www.dlib.org، ثم انسخ الملف المسمى dlib.html من خادم الويب إلى المتصفح، ثم أنهِ الاتصال".

أنواع مايم MIME types (١) :

مـن المعـروف أن أحـد ملـف البيانـات الموجـود فـي حاسـب آلـي معين هو ببساطة عبارة عن مجموعة من البتات Bits، ولكي يكون هذا الملف ذا قيمة فـإن هذه البتات يجب أن تكون موضحة أو يمكن تفسيرها. ففي المثال السابق ولكـي يمكـن عـرض الملـف المـسمى "dlib.html" بـشكل صـحيح، يجـب أن

(١) ملحقات بريد الإنترنت متعددة الأغراض Multipurpose Internet Mail Extensions .

يعرف المتصفح أن هذا الملف مكتوب بصيغة لغة تهيئة النص الفائق HTML، وتعتمد عملية التفسير هذه على نوع البيانات التي يتضمنها هذا الملف. وهناك نوعان شائعان من أنواع البيانات، هما: بيانات اتش تي إم إل HTML، وتستخدم مع ملف خاص بنص تم تهيئته عن طريق لغة تهيئة النص الفائق، وبيانات جي بيج Jpeg (١) وتستخدم مع ملف يعرض صورة تم بناؤه بصيغة Jpeg.

ويتم تحديد نوع البيانات سواء في الويب أو في العديد من تطبيقات الإنترنت المتنوعة من طريق خطة أو برنامج يعرف بمايم MIME - الاسم الحقيقي هو أنواع وسائط الإنترنت Internet Media Types، وقد تم تطوير برنامج مايم هذا بهدف وصف المعلومات التي ترسل عن طريق البريد الإلكتروني، ويستخدم في ذلك عملية ترميز ذات شقين Two Part Encoding: شق مع الجزء العام، وشق آخر مع الجزء الخاص. فعلى سبيل المثال يعد الترميز Text / ASCII هو نوع المايم المستخدم مع النص الذي تم ترميزه باستخدام نظام آسكي للترميز ASCII، أما Image / Jpeg هو نوع المايم المستخدم مع الصورة في صيغة Jpeg، و text / html تعبير عن نص تم تهيئته عن طريق لغة تهيئة النص الفائق، وهناك مجموعة قياسية من أنواع مايم تستخدم من قبل برامج حاسبات متعددة، كما يمكن وصف أنواع بيانات إضافية باستخدام تيجان تجريبية.

---

(٢) اختصار لـ Joint photographic expert group المجموعة التعاونية لخبراء التصوير الفوتوغرافي (المترجمان) .

وتكمن أهمية أنواع مايم على الويب في أن البيانات المرسلة من خلال أمر "احصل على" الخاص ببروتوكول نقل الملفات الفائقة Http Get Command يكون مرتبطاً معه نوع بيانات مايم، فعلى سبيل المثال الملف "dlib.org" يتضمن ما يشير إلى احتوائه على نوع من بيانات مايم، وهو نص مهيأ بلغة تهيئة النصوص الفائقة. وعندما يتلقى المتصفح ملفاً من هذا النوع، فإنه يعرف أن الطريقة المناسبة للتعامل مع هذا الملف هي ترجمته كنص مهيأ بلغة تهيئة النصوص الفائقة، ومن ثم عرضه على الشاشة.

إن كثيراً من أنظمة الحاسب الآلي تستخدم أسماء الملفات بطريقة غير مستساغة لتسجيل أنواع البيانات، فعلى سبيل المثال، تستخدم بعض برامج النوافذ أسماء ملفات تنتهي بالاختصارhtm بالنسبة لملفات بيانات مهيأة بلغة تهيئة النصوص الفائقة html، في حين تستخدم حاسبات تعمل بنظام يونكس Unix الاختصار html للغرض نفسه. وبذلك تعد أنواع مايم طريقة مرنة ومنتظمة جدًا لتسجيل البيانات النمطية وإرسالها Typed Data.

المعلومات المتاحة على الويب :

يمكن استخدام المكونات التي وصفت في الفقرات السابقة لإنشاء كثير من التطبيقات، غير أن بعضاً من هذه التطبيقات هو ما يمكن اعتباره مكتبات رقمية، أعني بذلك مجموعات منظمة من المعلومات المتنوعة، ولدينا هنا بعض النماذج لتطبيقات أخرى، فهناك شركات كثيرة لكل منها موقعها على الويب، وهذا الموقع يتضمن معلومات عن الشركة، وعن منتجاتها وخدماتها، كذلك يمكن شراء تذاكر الطيران من خادم الويب، كما أن هناك جمعيات لها مواقع على الويب تقدم من خلالها المعلومات لأعضائها،

وفضلاً عن ذلك هناك أشخاص لهم مواقعهم الخاصة على الويب أو لهم صفحاتهم الرئيسية الشخصية، كما أن نتائج البحوث العلمية يمكن أن تجد سبيلها قبل النشر على الويب. وبالرغم من ذلك هناك على النقيض الكثير من مواقع الويب التي لا يمكن وصفها إلا بسلة المهملات. وإجمالاً يمكن القول بأن المواقع المتنوعة تلك لها أهميتها في تطور المكتبات الرقمية، حيث إنها مسؤولة عن التطورات الفنية التي كانت وراء ما يمكن اعتباره قوالب البناء البسيطة التي تم وصفها في فقرات سابقة، كما أنها كانت مسؤولة كذلك عن عمليات التجريب، وعن وضع اتفاقيات تنظيم المواد، وعن الجودة العالية المتزايدة لأساليب التصميم الجرافيكي graphical design، وعن ظهور الأعداد الكبيرة من المنشئين المهرة، وغيرهم من المستفيدين ومديري مواقع الويب Webmasters المهرة كذلك. ومن الملاحظ أن حجم الويب قد دفع العديد من الشركات إلى تطوير منتجات يستخدم كثير منها الآن في بناء المكتبات الرقمية. ومما يؤسف له أن نجاح جميع مواقع الويب هذه غالباً ما يثقل أركان الإنترنت بما فوق طاقتها، بل إن هذه النجاحات قد خلفت من ناحية أخرى هموماً اجتماعية وقانونية تجاه السلوكيات غير السوية التي تمارس عبر الإنترنت.

اتفاقيات الويب Conventions of the Web :

من المعروف أن مواقع الويب الأولى أنشأها أفراد، وقد استندوا في ذلك إلى اعتبارات معينة كان النظر إلى المعلومات المناسبة يتم بموجبها، ثم تلا ذلك - بعد فترة وجيزة - ظهور الاتفاقيات التي تسعى إلى تنظيم المواد على تلك المواقع. وقد أتاحت الروابط الفائقة الفرصة لظهور مجموعة متنوعة

من اتفاقيات إتاحة المعلومات على مواقع الويب؛ حيث يمكن للمستفيدين أن يبحروا بفاعلية عبر بنيات مألوفة لهم. وبناء على ذلك فإن الاتفاقيات تكتسب أهمية بالغة في تصميم المكتبات الرقمية الموجودة على الويب. وبالرغم من أن تلك الاتفاقيات لم تأخذ طريقها مطلقاً عبر هيئات التوحيد القياسي، فإن تبنيها على نطاق واسع يضيف نوعاً ما من الترابط للويب، وهو الترابط الذي لم يكن موجوداً في سياق التقنية الأساسية.

اللوحة رقم (٢– ٧)

اتحاد الويب

The World Wide Web Consortium

بالرغم من عدم وجود جهة مركزية تضبط مسار الويب، فإنه يجب أن يكون هناك اتفاق على البروتوكولات الأساسية والصيغ والممارسات حتى يمكن للحاسبات المستقلة أن تعمل مع غيرها من الحاسبات الأخرى.

واعترافاً بهذه الحاجة فقد قام معهد ماساشوستس للتقنية MIT في عام ١٩٩٤م بإنشاء اتحاد الويب أو النسيج العنكبوتي العالمي (الذي يعرف بـ W3C)، وتم تعيين بيرنز- لي Tim Berners - Lee صاحب فكرة إنشاء الويب مديراً لهذا الاتحاد، ثم رأى المعهد بعد ذلك ضم شركاء دوليين إلى عضوية ذلك الاتحاد، وهم: المعهد الوطني لبحوث المعلومات والإلكترونيات في فرنسا Institute National de Recherche Informatique et en Autonalique ، وجامعة كيو في اليابان Keio University,s Shonan Fnjisawa Campus. ويتم تمويل اتحاد النسيج العنكبوتي العالمي من قبل المنظمات الأعضاء التي تضم معظم الشركات الكبرى التي تقوم بتطوير متصفحات الويب وخدمها

وغيرها من المنتجات ذات العلاقة.

ويعد اتحاد نسيج الشبكة العنكبوتية W3C منتدى محايداً، حيث تعمل المنظمات فيه مع بعضها البعض بمواصفات مشتركة من أجل خدمة أهداف الويب، وذلك عبر سلسلة من المؤتمرات وورش العمل وعمليات التصميم، كما أنه يوفر مجموعة من المعلومات عن الويب لكل من المطورين والمستفيدين، بما في ذلك عينات أو نماذج الرموز samples of code التي تساعد على نشر المواصفات القياسية. كما أن هذا الاتحاد يعمل بشكل كبير مع لجنة هندسة للإنترنت من أجل نشر معايير للتقنية الأساسية للويب مثل بروتوكول نقل الملفات الفائقة HTTP، ولغة تهيئة النصوص الفائقة، HTML والمحددات الموحدة لمواقع المصادر URLs.

وبحكم عمل اتحاد نسيج الشبكة العنكبوتية كجهة محايدة في مجال تهيمن عليه الشركات المتنافسة، فإن سلطته تعتمد على مدى تأثيره، ومن أبرز نجاحاته في هذا المجال تطويره السريع لأحد المعايير الصناعية المستخدمة في تقدير المحتوى أو المضمون Content، والذي يعرف بمعيار بيكس PICS، الذي طور استجابة للمخاوف السياسية التي كانت سائدة في الولايات المتحدة بشأن قدرة بعض الشركات الصغرى على إتاحة الصور الخليعة وبعض المحتويات الأخرى غير المرغوب فيها. ومما يذكر أن اتحاد نسيج الشبكة العنكبوتية نشط نشاطاً كبيراً في السنوات الأخيرة في تطوير اللغة

الموسعة لتهيئة النصوص Extensible Mark Up Language (XML) [(١)].

وبرغم مما قد يبدو لبعض الشركات مثل: مايكروسوفت، ونتسكيب، وغيرهما، من قناعة بأنها قد تجني أرباحها عن طريق تصدير المنتجات التي لا تتوافر لها الخصائص القياسية، فإن مثل هذه الخصائص تضر بتجانس الويب أو توافقه. وبهذه المناسبة فالشكر والثناء مكفولان لاتحاد نسيج الشبكة العنكبوتية على الأسلوب المسؤول الثابت الذي تواصل به تقنية الويب مسيرتها نحو التطور.

مواقع الويب Web Sites :

موقع الويب هو مجموعة من المعلومات التي يستشعر المستفيد بأنها وحدة واحدة، وغالباً ما يكون موقع الويب مرتبطاً بخادم ويب وحيد، ومع ذلك فمن الممكن أن يُحمل أحد المواقع الضخمة على أكثر من خادم واحد، كما يمكن أن يستضيف خادم واحد أكثر من موقع واحد من مواقع الويب.

***الصفحات الرئيسية* Home Pages :**

الصفحة الرئيسية هي الصفحة الأولى أو التمهيدية لمجموعة من معلومات الويب، وعادة ما يكون لكل موقع على الويب صفحته الرئيسية، وإذا لم يتمكن العنوان الموجود على المحدد الموحد لموقع المصدر URL من تحديد اسم أحد الملفات، فإن الخادم يقوم ـ وبشكل تقليدي ـ بتوفير صفحة تسمى "بكشاف ملفات لغة تهيئة النصوص الفائقة Index.html"؛ ومن ثم

---

(١) إصدارة من SGML خاصة بالوثائق المعروضة على الويب تساعد المستخدمين على إضافة مهام للوثيقة غير موجودة في HTML (المترجمان) .

فـإن الـمحدد الموحـد الموسـوم

بـ " http://www.loc.gov/" يترجم على النحو التالي :

http://www.loc.gov/index.html

ومـن الملاحـظ أن لكـل مُـصمم أفكـارًا مخـتلفة عـن كيفـية تنظـيم الـصفحة الرئيـسية، ومـع ذلـك، وكمـا هـو الحـال بالنـسبة لـصفحات العـناوين فـي الكـتب المطـبوعة والتـي تتـبع فـي تنظـيمها بعـض الأعـراف المحـددة، فـإن الـصفحة الرئيسية عـادة مـا تعطـي فكـرة عامـة عن الموقـع، وعـادة مـا تكـون فـي أحسن حالاتهـا معلـومات عن الموقع، وقائمـة بمحتوياته، بالإضـافة إلـى بعض وسائل المساعدة لإيجاد المعلومات.

ويمكـن لمجموعات صـغيرة مـن المعلـومات داخـل موقع معين أن تسمى كذلك بالصفحات الرئيسية، وهو أمر شائع بالنسبة للمعلومات المتصلة بأقسام أو مـشروعات أو خـدمات محـددة لهـا صـفحاتها الرئيـسية داخـل أحـد مواقـع الويب الخاص بإحدى الشركات أو المؤسسات.

الأزرار Buttons :

تحـرص معظـم صـفحات الـويب الرئيـسية علـى توفيـر أزرار أو أيقونـات لمـساعدة المـستفيدين علـى الـتجول فـي الموقـع، ومـن أمثلـتها: الـزرhome - الخـاص بالعـودة للـصفحة الرئيـسية، والـزرnext - الخـاص بالانـتقال للـصفحة التالـية، والزر Previous - الخاص بالانتقال إلى الصفحة السابقة، وجميع هذه الأزرار روابط فائقة للأجزاء الأخرى بالموقع.

التنظيم الهرمي Hierarchical Organization :

تنظم العديد من مواقع الويب من وجهة نظر المستفيد بأسلوب هرمي، وعبر الروابط المتوافرة بالصفحة الرئيسية Home يتم الوصول إلى أقسام رئيسية قليلة؛ ثم تقوم هذه بدورها بالوصول إلى معلومات أكثر تحديداً... وهكذا. وغالباً ما تسمح الأزرار الموجودة في كل صفحة للمستفيد بالرجوع ثانية إلى المستوى الأعلى السابق في المستوى الهرمي، أو تسمح له بالانتقال إلى الصفحة التالية على المستوى نفسه.

الويب مكتبة رقمية The Web as a Digital Library :

كثيراً ما يتحدث بعض الناس عن تقنية الويب كما لو كانت بديلاً مؤقتاً دون المستوى يتخذ إلى حين إنشاء المكتبات الرقمية، ويرجع سبب هذا التوجه الفكري إلى صعوبة تقبل كثير من المهنيين القول بأن قصب السبق في مجال المكتبات الرقمية كان من نصيب الفيزيائيين في أحد مختبرات سويسرا، ولم يكن من نصيب المكتبيين المعروفين، أو من علماء الحاسبات الآلية. ولكن ما يمكن قوله إن الويب ليست تحويلة مؤقتة يعتمد عليها لحين ظهور المكتبات الرقمية، وإنما هي المركز الذي تدور حوله كل جهود تطوير المكتبات الرقمية.

وبالإضافة إلى ما سبق، ينتقص بعضٌ من غير المتآلفين مع المجموعات المباشرة أو المتاحة على الخط المباشر On – line Collection من قدر المعلومات المتاحة على الويب، ومن أكثر دعاواهم في هذا الصدد دعوتان: الأولى أن تلك المعلومات لا تتوافر لها ضمانات الدقة؛ والثانية أن البحث عن تلك المعلومات أمر يكاد يكون من ضروب المستحيل. وإذا سلمنا بصحة بعض ما جاء في هاتين الدعوتين، يمكن القول بأنهما أبعد ما تكونان عن

الحقيقة الكاملة؛ فهناك كم هائل من المعلومات المتاحة على الويب، وإن كانت هناك معلومات ذات قيمة متواضعة فإن هناك كثيراً من خدمات الويب تلقى عناية كبيرة، ويتوافر بها معلومات على قدر كبير من الجودة، كذلك يمكن القول بأن الوصول للمعلومات على الويب قد يواجه بعض الصعوبات، ولكن هناك في المقابل أدوات ومرافق بحثية تمكن المستفيد- وبقدر قليل من الجهد- من استكشاف تلك المعلومات المتوافرة والوصول إليها.

إن الويب التي نعيشها اليوم هي بداية الطريق وليست نهايته، وإذا كانت الافتراضات المبسطة لما وراء التقنية تبدو براقة، فإن لها محدودياتها. وبالرغم من أن الويب التي نعيشها اليوم تقدم أسساً لبناء مكتبات رقمية بمجموعات جيدة، وبخدمات متطورة، وتقنيات راسخة، فإن كثيراً من البحوث الجارية الآن في مجال المكتبات الرقمية ينظر إليها وكأنها تسعى إلى توسعة دعائم البناء الأساسية للويب. ويمكننا أن نتوقع في حدود خمسة وعشرين عاماً من الآن أن تكون المكتبات الرقمية مختلفة تماماً عما هي عليه الآن، وفيها سيصعب علينا تذكر ما كانت عليه في الأيام الأولى من تاريخها، وقد تصبح كلمات مثل "الإنترنت" و"الويب" تاريخاً، أو ربما تنطبق على النظم التي سيكون من الصعب التعرف عليها، وأنها ستكون خلفاً لنظم ليست معروفة أنها حفيدة لنظم سابقة، وسوف تتمكن المكتبات الرقمية من استيعاب مواد وتقنيات من أماكن مختلفة، أما للسنوات القليلة القادمة فيمكن أن نتوقع رؤية الإنترنت والويب بمثابة الأسس التي ستبنى عليها مكتبات المستقبل، وكما حدث بالفعل عندما تطورت البرمجية المعقدة في الحاسبات الشخصية المبكرة إلى نظم تشغيل حديثة، فإن الويب يمكن أن

تصبح أساساً لأجيال عديدة من المكتبات الرقمية.

الفصل الثالث

المكتبات والناشرون

حازت المكتبات عظيم الشرف بسبب تبنيها المبكر للتقنيات الحديثة، كتقنيات الوسائط الميكروفيلمية، وخدمات المعلومات المتاحة على الخط المباشر، والأسطوانات المدمجة CD-ROMs (وتعد ثاني آلة طباعة هي تلك التي صُدّرت للمكتبة الوطنية في البرازيل بغرض كتابة بطاقات فهرس تلك المكتبة). وتعد الإنترنت والويب أحدث نموذجين لمدى احتضان المكتبات للتقنيات الحديثة، ولمدى تكييف تلك التقنيات لاحتياجاتها.

ومن الملاحظ أن المنتجات التجارية السائدة لا تتفق دائما واحتياجات المكتبات؛ حيث تستهدف صناعة الحاسبات بمنتجاتها الأسواق الكبرى ذات الطابع التجاري والعلمي والطبي والعسكري، وقد مولت وزارة الدفاع الأمريكية وبعض الشركات الأخرى كشركة أي بي إم IBM الغالبية العظمى من البحوث الأساسية التي كانت وراء ظهور تلك المنتجات، وحتى وقت قريب لم تكن تلك الأسواق الكبرى تولي إدارة المعلومات المتوافرة بالمكتبات الرقمية اهتماماً ملموساً، كما لم تكن تلك المنتجات التي تطرحها صناعة الحاسبات تلبي الاحتياجات الضرورية لإدارة المعلومات. وبناء على ذلك فقد اعتادت المكتبات الحصول على التقنيات الأساسية من مجالات أخرى ثم توظيفها وفق احتياجاتها.

تقاسم الموارد والفهارس المباشرة :

لكلمة المشابكة networking في مجال المكتبات معنيان؛ فقبل انتشار شبكات الحاسبات التي نعرفها اليوم، كان للمكتبات أنشطة متنوعة لتقاسم مواردها، ومن هذه الأنشطة تبادل الإعارة بين المكتبات، وامتيازات القراءة (١) المتبادلة بين المكتبات، وتبادل النسخ المصورة. وقد كان يعبر عن كل تلك البرامج إما بمصطلح "المشابكة networking" أو "تقاسم الموارد resource sharing". وقد عززت خدمات تقاسم الموارد كخدمات توفير النسخ المصورة وتوصيل الوثائق التي قدمتها المكتبة البريطانية مبدأ "إتاحة المعلومات availability of information" حول العالم. ومع ذلك لم يكن لشيء في تقاسم الموارد من أثر ينافسه كما كان لشبكات الحاسب الآلي.

ولجميع المكتبات فهارسها التي تشتمل على سجلات (٢) تعرف بمجموعاتها من مصادر المعلومات، وعادة ما تتكفل هذه الفهارس بمساعدة المستفيدين في البحث عن مقتنيات المكتبة، ثم تقديم البيانات الببليوجرافية التي تصف تلك المقتنيات، وهي بالإضافة إلى ذلك تعد من الأدوات الهامة التي يستعان بها في إدارة المقتنيات.

ويعد مجال الفهرسة من المجالات التي يستخدم فيها المكتبيون مصطلحات دقيقة قد يكون بعضها غير مألوف لغيرهم من خارج المجال، فبينما يستخدم غير المتخصصين مصطلح "فهرس" كمصطلح عام، فإنه يكتسب معنى محدداً عند المكتبيين، وهو "مجموعة من التسجيلات الببليوجرافية التي أعدت حسب

---

(١) لعل المؤلف يقصد الإعارة وليس القراءة (المترجمان).

(٢) جرى عرف المكتبيين على تسميتها ببطاقات الفهرس إذا كان مطبوعاً وبالتسجيلات إذا كان الفهرس آلياً (المترجمان) .

قواعد مقننة"، كذلك يستخدم المكتبيون مصطلح "المنفردات أو الكتب أحادية الموضوع monographs" بديلاً عن المصطلح المتداول يومياً وهو "الكتاب Book". ومع مرور الأيام تم تبادل المعلومات التي تتضمنها تسجيلات فهارس المنفردات التي تم إعدادها وفق قواعد الفهرسة، وتستخدم الدول الناطقة بالإنجليزية (١) قواعد أنجلو الأمريكية للفهرسة AACR، والتي تعرف الطبعة الثانية منها اختصاراً بـ AACR2.

ولا شك أن فهرسة أي عمل منفرد تستغرق وقتاً طويلاً، كما تتطلب خبرة بهذه العملية، وتوفيراً للتكاليف فقد سعت المكتبات إلى الفهرسة التعاونية (٢)، وتعمد المكتبات الكبرى التي تضطلع بفهرسة أعداد ضخمة من المنفردات كمكتبة الكونجرس ومكتبات الجامعات الكبرى إلى إتاحة سجلات فهارسها إلى المكتبات الأخرى مجاناً، وبدلاً من أن تقوم تلك المكتبات بتكرار فهرسة الكتاب الذي تقتنيه حديثاً، تكتفي فقط بفهرسة الكتب التي لا تجد لها سجلات متاحة في فهارس تلك المكتبات الأخرى، وذلك بالطبع بعد البحث عنها في فهارسها (٣).

ومنذ أواخر الستينات الميلادية من القرن العشرين بدأت عملية الفهرسة التعاونية تأخذ أبعاداً أخرى؛ حيث تعتمد على الحاسبات الآلية، وتعد الأداة الأساس من الناحية الفنية لعرض هذه السجلات هي صيغة "الفهرسة المقروءة

---

(١) بحكم بعد المؤلف عن مجال المكتبات والمعلومات لم تتوافر لـه المعلومات الكافية عن مدى استخدام هذه القواعد على النطاق الدولي (المترجمان).

(٢) استخدم المؤلف مصطلح تقاسم سجلات الفهارس share catalog records (المترجمان) .

(٣) تعرف هذه العملية في عرف المكتبيين بالفهرسة المنقولة (المترجمان) .

آلياً Machine Readable Cataloging، المعروفة اختصاراً بمارك MARC"، التي تبنت تطويرها هنريت أفرام Herniette Avram وزملاؤها في مكتبة الكونجرس [1]. وقد كانت هذه الصيغة في أول الأمر صيغة لتوزيع سجلات الفهارس على شرائط ممغنطة. ومن الجدير بالملاحظة أن مصطلح "الفهرسة المقروءة آلياً باستخدام صيغة مارك MARC Cataloging"، كثيرًا ما يستخدم على مستوى الممارسة العملية بمعناه العام للدلالة على كل من سجلات الفهارس والصيغة الآلية التي تختزن بها تلك السجلات. ولم تعد الفهرسة المقروءة آلياً قاصرة على المنفردات، بل اتسع مجال تطبيقها لتشمل الآن جميع أنواع مقتنيات المكتبات كالمسلسلات والمواد الأرشيفية والمخطوطات. وتوضح اللوحة (٣-١) أحد نماذج صيغة مارك للفهرسة الآلية.

وقد أدى تطوير صيغة مارك للفهرسة الآلية إلى نوعين من النظم المبنية على الحاسبات الآلية في مجال المكتبات، يتمثل أول هذين النوعين في عملية "الفهرسة التعاونية shared cataloging" التي يرجع فضل الريادة فيها إلى فريد كيلجور Fred kilgour (الذي أنشأ مركز الحاسب الآلي للمكتبات على الخط المباشر المعروف بـ أو سي إل سي "OCLC" في عام

(٤) بدأت هنريت أفرام مشروع التحسيب في مكتبة الكونجرس في أواسط الستينات (المترجمان).

١٩٦٧م)[١]، حيث يتوافر لهذا المركز نظام آلي ضخم وصل عدد السجلات التي يضمها إلى أكثر من ٣٥ مليون سجل بصيغة مارك [٢]، بما فيها السجلات التي يتلقاها المركز من مكتبة الكونجرس. وعندما تقوم أي مكتبة من المكتبات الأعضاء في هذا المركز باقتناء أحد الكتب وترغب في فهرسته، فإنها تقوم أولاً بالبحث عنه في قاعدة بيانات هذا المركز، فإذا وجدت سجل هذا الكتاب فإنها تحمله على نظام الحاسب الآلي الخاص بها، وتقوم في المقابل بفهرسة الكتاب وتحميل سجله في قاعدة بيانات هذا المركز إذا لم تجد له سجلاً مسبقًا، وتعرف هذه العملية "بالفهرسة المنقولة Copy cataloging". وبموجب عملية الفهرسة المنقولة هذه يتم فهرسة العنوان الواحد مرة واحدة، ومن ثم فإن هذا الجهد الفكري المبذول في عملية الفهرسة يكون قد تقاسمته جميع المكتبات. وقد أدى نجاح عملية الفهرسة المقروءة آلياً MARC وجهود مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC، إلى دفع خدمات مكتبية أخرى مشابهة حول العالم إلى محاكاة هذا المركز.

وقد شجعت إتاحة سجلات مارك على ظهور تطور آخر تمثل في أن المكتبات بدأت في إنشاء فهارس مباشرة مستقلة لمقتنياتها، وكثيرًا ما يتم

---

(١) كان هذا المركز يعرف عند إنشائه بمركز مكتبات الكليات بأوهايو، Ohaio College e Library Center, ثم يعرف الآن مركز الحاسب الآلي للمكتبات على الخط المباشر Online Computer Library Center (المترجمان) .

(٢) احتفل في عام ٢٠٠٢م بإضافة التسجيلة رقم خمسين مليونًا لقاعدة بيانات هذا المركز (المترجمان).

الاعتماد في الحصول على الجزء الأكبر من تلك الفهارس على عملية "الفهرسة المنقولة"، التي أشرنا إليها في الفقرة السابقة. واليوم يمكن القول بأن لكل مكتبة كبرى تقريباً في الولايات المتحدة الأمريكية فهرسها المباشر الذي يعرف في اصطلاح المكتبيين "بالفهرس المتاح للجمهور على الخط المباشر واختصارًا " أوباك" Online Public Access Catalog (OPAC)"، وقد سعت مكتبات كثيرة إلى بذل جهود عظيمة لتحويل فهارسها البطاقية القديمة إلى فهارس آلية بصيغة "مارك"، ولذلك يصبح الفهرس المتاح على الخط المباشر هو السجل الشامل لمجموعاتها كافة، وقد أدى هذا في حد ذاته إلى تجنب إنشاء فهرس مباشر للمقتنيات الحديثة وفهرس بطاقي آخر مستقل للمقتنيات القديمة، وقد أكملت جامعة هارفرد Harvard University في الفترة الأخيرة عملية تحويل كاملة لما يقرب من خمسة ملايين بطاقة، وبتكلفة تقدر بحوالي خمسة عشر مليون دولار أمريكي.

إن صيغة مارك كانت صيغة مبتكرة في وقت كانت فيه معظم نظم الحاسبات الآلية تعرض النصوص كحقول ثابتة الطول Fixed-length field بحروف كبيرة فقط، وقد بقيت هذه الصيغة هي الصيغة الحيوية للمكتبات بالرغم من تقادمها.كما أنه ليس من السهل التكهن بمستقبلها في ظل الاستثمار المكثف الذي توليه لها المكتبات، وأياً كان هذا المستقبل، فإن صيغة مارك كانت إنجازاً ريادياً في تاريخ كل من استخدام الحاسبات والمكتبات على حد سواء، وهي الصيغة الأساسية التي يجب أن تستوعبها

المكتبات الرقمية [1].

اللوحة رقم (٣ - ١)

سجلات مارك

هذه تسجيلة ببليوجرافية لأحد المنفردات، كتبت على النحو التالي:

Caroline R. Arms, editor, Campus strategies for libraries and electronic information. Bedford, Mass: Digital Press, 1990.

ومن خلال البحث في فهرس مكتبة الكونجرس عن طريق واجهة تعامل إحدى الطرفيات، تم عرض مدخل هذا العمل المشار إليه سابقاً في شكل تظهر فيه المعلومات في صيغة مارك على النحو التالي:

89-16879r93 &001

Z675.U5C161990 &050

027.7/097320 &082

Campus strategies for libraries and electronic &245

information/Caroline Arms, editor.

{Bedford, Mass.}:Digital Press,c1990. &260

xi,404p.:ill.;24cm. &300

&440 EDUCOM strategies series on information technology

bibliographical references(p.{373}-381). Includes &504

ISBN1-55558-036-X:$34.95 &020

(١) ظهرت الإصدارة الأخيرة من مارك بعنوان مارك ٢١ لتتواءم مع كثير من التغيرات التي تظهر خلال القرن الحادي والعشرين، وخاصة ما يتصل بمصادر المعلومات الإلكترونية وغيرها (المترجمان).

Academic libraries—United States--Automation. &650

Libraries and electronic publishing—United States. &650

Library information networks—United States. &650

Information technology—United States. &650

Arms ,Caroline R.(Caroline Ruth) &700

DLCDLCDLC&043n-us--- &040

CIPver.br02toSL02-26-90 &955

APIF/MIG &985

ويلاحظ أن البيانات مقسمة إلى حقول، وكل حقل منها يرمز له برمز مكون من ثلاثة أرقام، فعلى سبيل المثال: الرقم 440 يمثل حقل عنوان السلسلة، والرقم 650 يمثل حقل رؤوس موضوعات مكتبة الكونجرس، وهناك قواعد معقدة تعرف المفهرس بالحقول التي يجب استخدامها، وكيف يمكن تفسير العلاقات بين عناصرها.

ومن الجدير بالإشارة هنا أن عملية الترميز الفعلية أعقد بكثير مما تبدو عليها الآن في النموذج المعروض؛ فصيغة مارك الكاملة تتكون من مجموعة محددة مسبقاً من الحقول، وكل حقل من هذه الحقول يحدد بتاج معين، مع ملاحظة أن الحقول الرئيسية تحدد بتاج مكون من ثلاثة أرقام، في حين تحدد الحقول الفرعية بحروف فردية.

ولإلقاء نظرة سريعة على كيفية ترميز المعلومات بهذه الصيغة، لاحظ الحقل رقم 260 الذي يبدأ بالتاج 260 &، ففي تسجيلة فعلية من تسجيلات مارك، يتم ترميز هذا الحقل على النحو التالي، حيث تشير التركيبة "abc"

والتي ترد تالية لتاج الحقل إلى وجود ثلاثة حقول فرعية:

&2600#abc#[Bedford,Mass.] #Digital Press, #c1990.%

فالحقل الفرعي الأول وهو حقل مكان النشر حدد بالتاج a، في حين حدد الحقل الفرعي الثاني وهو اسم الناشر بالتاج b،كما حدد الحقل الفرعي الثالث الخاص بتاريخ النشر بالتاج c.

ربط الفهارس المباشرة والمعيار Z39.50 :

مع ثمانينات القرن العشرين الميلادي بدأت المكتبات الجامعية في التوجه نحو ربط فهارسها المتاحة على الخط المباشر بالشبكات، ومع حلول عام ١٩٨٤م، وبوجود شبكة شاملة بالمدينة الجامعية campus network لكلية دارتماوث Dartmouth college، وحيث تم ربط الحاسب الآلي الذي يُحمّل عليه فهرس هذه المكتبة بتلك الشبكة، فإنه يمكن لجميع من يعملون بتلك الكلية وتتوافر لهم طرفيات أو حاسبات شخصية متصلة بالشبكة أن يبحثوا في فهرس المكتبة. كما يمكن لجميع منسوبي الجامعة أن يستخدموا حاسباتهم في بحث فهارس مكتبات الجامعات الأخرى. وعندما تم ربط شبكة تلك الجامعة بالإنترنت كان في مقدور الجميع على مستوى العالم البحث في فهرس مكتبتها. بناءً على ذلك يمكن القول بأن عملية تقاسم [١] فهارس المكتبات كانت أول أشكال المشاركة التعاونية الواسعة النطاق للمعلومات المتاحة عبر الإنترنت.

وفي أواخر سبعينات القرن العشرين الميلادي كانت العديد من المرافق الببليوجرافية بما فيها مكتبة الكونجرس ومجموعة مكتبات البحث وشبكة

(١) كان من الأجدر بالمؤلف تسميتها مشابكة وليس تقاسمًا (المترجمان).

معلومات مكتبات واشنطن قد بدأت في تدشين مشروع عرف "بمشروع النظم المترابطة Linked systems project" الذي طور بروتوكول يعرف الآن بـ Z39.50. يهدف إلى إتاحة الفرصة لأي حاسب أن يبحث عن المعلومات المتوافرة في حاسب آخر، وقد استخدم هذا البروتوكول أساسًا في البحث عن سجلات مارك، غير أنه كان مرناً ولم يتقيد بنظام مارك، ومن الناحية الفنية يحدد بروتوكول Z39.50 القواعد التي تسمح لأحد الحاسبات بالبحث في قواعد بيانات حاسب آخر واسترجاع سجلاته، ويستخدم هذا البروتوكول الآن على نطاق واسع في عمليات التشغيل المتداخل Interoperability بين مجموعة مختلفة من نظم الحاسب الآلي.

المستخلصات والكشافات :

إذا كانت فهارس المكتبات هي مصادر معلومات أساسية (١) عن المنفردات، فإنها ليست بهذه الدرجة بالنسبة للدوريات؛ فالفهارس لا تقدم سوى سجل مختصر لدورية معينة بكامل إصدارتها، ولأن هذا السجل ليس له تلك القيمة الكبرى التي تمكن من يبحث عن مقالة معينة من الوصول إليها؛ من ثم برز الدافع لإيجاد خدمات تكشيف واستخلاص تسعى إلى مساعدة الباحثين في الوصول إلى مقالات محددة، ومن نماذج تلك الخدمات قاعدة مدلاين Medline (للإنتاج الفكري الطبي)، وقاعدة المستخلصات الكيميائية Chemical Abstract (للإنتاج الفكري في مجال الكيمياء) وقاعدة إنسبك Inspec (للإنتاج الفكري في مجال الفيزياء والحاسبات). ومع أن هذه

(٢) احتكاماً لعرف المكتبيين فهي مصادر معلومات من الدرجة الثالثة على اعتبار أن ثمة مصادر معلومات أولية وثانوية (المترجمان).

الخدمات تتفق في بنيتها الأساسية، فإن ثمة فروقاً بينها في كثير من التفاصيل، فالمهنيون [١] الذين لديهم المعرفة الكافية بمجال موضوعي معين يقومون بقراءة كل المقالات في أكبر عدد من دوريات هذا المجال، ثم يقومون بتحديد المصطلحات التكشيفية index terms أو يقومون بصياغة المستخلصات. ومن الملاحظ أن هذه الخدمات تستخدم مصطلحات تكشيفية مأخوذة بعناية كبيرة من إحدى لغات التكشيف المقيدة، مثل قائمة رؤوس الموضوعات الطبية المعروفة اختصاراً بـ "مش MeSh " والتي تستخدمها المكتبة الوطنية للطب في إعداد بيانات قاعدة معلوماتها المعروفة بالمدلاين، هذا في الوقت الذي لا تسير فيه الأمور في بعض الخدمات بالقدر نفسه من الإحكام، كما أن هناك بعض خدمات الاستخلاص تضطلع هي بنفسها بمهام إعداد مستخلصاتها، في حين يستخدم البعض الآخر مثل إنسبيك المستخلصات التي يعدها ناشرو الدوريات التي يقومون بإصدارها.

لقد بدأت معظم تلك الخدمات بإنتاج مجلداتها المطبوعة بهدف بيعها للمكتبات، غير أن الموقف تغير اليوم؛ حيث تتم الغالبية العظمى من عمليات بحث هذه الخدمات باستخدام الحاسبات (يرجع تاريخ البحث الآلي في الكشافات إلى مرحلة ما كان يعرف بالتجهيز على دفعات Batch processing، واستخدام الشرائط الممغنطة في عمليات التخزين). وتقوم بعض خدمات التكشيف بتشغيل نظم آلية تتيح عمليات البحث فيها مقابل رسوم محددة، في حين تقوم جهات أخرى [٢] بإتاحة بياناتها لطرف ثالث [١] يقوم بدوره بتقديم

(١) يقصد بالمهنيين المكشفين ومن يضطلع بعمليات الاستخلاص (المترجمان) .

(١) عادة ما تعرف هذه الجهات بمنتجي قواعد البيانات (المترجمان).

خدمات البحث في تلك البيانات على الخط المباشر. كما تقوم العديد من المكتبات الكبرى بإتاحة فهارسها على حاسباتها الخاصة بها، أضف إلى ذلك أن كثيراً من تلك الفهارس تتاح الآن على الأقراص المدمجة CD-ROMs.

وما أن أصبحت فهارس المكتبات متاحة على الخط المباشر، حتى بدأت المكتبات في إتاحة بيانات أخرى لمواد أخرى مثل مستخلصات المقالات والكشافات والأعمال المرجعية. ويمكن أن تختزن مصادر المعلومات هذه في حاسب مركزي بحيث يتم عرض السجلات المختزنة على النهايات الطرفية والحاسبات الشخصية. وتعد الأعمال المرجعية التي تتكون من مداخل مختصرة مناسبة جداً بشكل خاص لهذا النمط من التوزيع لأن المستفيدين ينتقلون بسرعة من مدخل إلى آخر، وسوف يتقبلون عرضاً يتضمن رموزًا نصية مع قدر بسيط من التوحيد في الصيغ. وتعد عمليات الاسترجاع السريع والبحث المرن أكثر أهمية من جماليات عرض المخرجات على الشاشات. وفيما يلي قائمة بنماذج فعلية لبعض مصادر المعلومات المتعددة التي أتاحتها جامعة كارنيجي ميلون Carnegie Mellon University على الخط المباشر خلال عام ١٩٩٨م :

- فهرس مكتبة جامعة كارنيجي ميلون.
- قائمة دوريات مكتبة جامعة كارنيجي ميلون.
- سجلات ببليوجرافية لصور ورسومات معمارية.
- دليل منسوبي جامعة كارنيجي ميلون Who's who at CMU

(٢) عادة ما تعرف هذه الجهات بموزعي قواعد البيانات أو مورديها (المترجمان) .

- معجم التراث الأمريكي.
- مستخلصات الدوريات.
- قاعدة معلومات ABI/inform (للإنتاج الفكري في مجال إدارة الأعمال).
- قاعدة معلومات INSPEC (في مجال الفيزياء والإلكترونيات وعلوم الحاسب).
- دليل البحوث Research Directory (لجامعة كارنيجي ميلون).

ومما تجدر الإشارة إليه أن بعضاً من هذه المجموعات المتاحة على الخط المباشر تقدم معلومات ذات طابع محلي أو داخلي مثل دليل أعلام جامعة كارنيجي ميلون (وهو دليل لأعضاء هيئة التدريس والطلاب والموظفين بالجامعة). نظراً لأن رواد المكتبة لا يهتمون فقط بالمواد المنشورة رسمياً أو أكاديمياً، وتقدم المكتبات العامة- بشكل خاص - أنواعاً متعددة من المعلومات تبدأ بنماذج دفع الرسوم وتنتهي بجداول مواعيد الحافلات، كذلك تسمح كشافات النصوص الكاملة ومتصفحات الويب للمواد المكتبية التقليدية وغير التقليدية للاندماج عبر نظام واحد يتم التعامل معه من خلال واجهة تعامل موحدة. وقد أصبح هذا الاتجاه مقنناً بشكل واضح لدرجة أن من الصعوبة تصور أنه كان من النادر قبل السنوات القليلة الماضية تحقيق دمج المعلومات المستخلصة من مصادر متنوعة.

لا شك أن إتاحة كميات كبيرة من المعلومات على الخط المباشر، والمحافظة على تحديث هذه المعلومات عملية باهظة التكاليف، وبالرغم من

الانخفاض المستمر لأسعار المكونات المادية للحاسبات Hardware، فإن ارتفاع هذه الأسعار لا يزال ملموساً. ومن أكثر تلك التكاليف الباهظة ما يرتبط منها بعملية إصدار تراخيص البيانات والترخيص للأشخاص الذين يتعاملون مع كل من الجوانب المالية وملفات البيانات الضخمة، ولتقليل حدة هذه التكاليف كونت المكتبات فيما بينها اتحاداً بحيث يمكن لمجموعة واحدة من البيانات المتاحة على الخط المباشر أن تقوم بخدمة مكتبات متعددة.

استرجاع المعلومات :

تعد عملية استرجاع المعلومات من الموضوعات المحورية للمكتبات؛ فالمستفيد الذي قد يكون عالماً أو طبيباً أو محامياً عادة ما يكون مهتماً بالمعلومات حول موضوع معين، ويرغب في معرفة ما هو متوافر في مجموعة المكتبة ويغطي هذا الموضوع، ويتطلب هذا الأمر بلا شك برمجية متخصصة، وقد بدأت المكتبات منذ منتصف الثمانينات الميلادية من القرن العشرين استخدام حاسبات تتوفر بها برمجيات حاسوبية قادرة على بحث النصوص الكاملة لمجموعات ضخمة. وكانت تسجيلات مارك لمقتنيات المكتبات أولى البيانات التي تم تحميلها على هذه الحاسبات، ثم تبعتها الأعمال المرجعية القياسية. وتعني عملية بحث النصوص الكاملة أن المستفيد يمكن أن يقوم بإجراء عملية البحث باستخدام أية كلمة توجد بالسجل، ولا يشترط في ذلك أن تتوافر للمستفيد المعرفة الكاملة ببناء تلك السجلات ولا بالقواعد التي تحكم إعدادها.

وبالرغم من أن البحث في هذا المجال يعود إلى نحو ثلاثين عامًا

مضت، فإن اتجاهاته لم تتغير بشكل كبير؛ فالمستفيد يعبر عن حاجته للمعلومات باستفسار يتقدم به، وهذا الاستفسار قد يرد في شكل كلمة واحدة (مثل كلمة زهرة القرنبيط Cauliflower)، أو جملة (مثل المكتبات الرقمية Digital Libraries) أو عبارة مطولة (مثل: في أي سنة خرج داروين في رحلة برية مصطحباً معه البيجل Beagle [١]). ومهمة الاسترجاع هي إيجاد المواد التي تضاهي هذا التساؤل في المجموعة، ونظراً لعدم توافر الوقت الكافي لأي حاسب للبحث في المجموعة بأكملها لكل تساؤل والتعامل مع كل معلومة بشكل مستقل،كان لزاماً أن تتوافر للحاسب كشافات من نوع معين [٢]، تمكنه من استرجاع المعلومات من خلال البحث في مداخلها.

وبقدر ما وصلت إليه الحاسبات من تطور، ومن انخفاض في أسعار وسائط التخزين، فقد انتقلت أساليب استرجاع المعلومات من عمليات البحث المحدود بالسجلات المختصرة (مثل سجلات الفهارس أو تلك السجلات المستخدمة من قبل خدمات التكشيف والاستخلاص) إلى عمليات بحث النصوص الكاملة عن أي كلمة في مجموعات ضخمة. ومنذ وقت مبكر دفعت تكلفة التخزين والتجهيز الآلي إلى التفكير في تطوير كل من أساليب التخزين المدمج أو المكتنز compact storage والخوارزميات الحاسوبية الفعالة. وحديثاً كثفت البرامج البحثية للويب اتجاهاتها البحثية نحو أساليب بحث المعلومات الكثيرة الموزعة عبر العديد من الحاسبات.

---

(١) البيجل : كلب صيد صغير القوام ناعم الوبر (المترجمان) .

(٢) عادة ما تعرف هذه الكشافات بالكشافات المصنفة inverted files (المترجمان) .

لغات التهيئة أو الترميز Markup Languages [١] :

يجمع المكتبات ودور النشر اهتمام مشترك يتمثل في استخدام الحاسبات الآلية في العرض الكامل والدقيق full richness للمواد النصية؛ لأن الوثائق النصية أكبر من مجرد سلسلة متتابعة من الحروف، بل قد تتضمن على سبيل المثال رموزاً رياضية أو موسيقية وحروفاً من جميع اللغات وتشكيلة من الحروف المطبعية Fonts، بالإضافة إلى عناصر بنائية كالعناوين الرئيسية والحواشي والكشافات. والطريقة المفضلة لاختزان وثيقة في حاسب معين هي ترميز الخصائص المصاحبة لنص الوثيقة، واختزانها مع النص والأشكال والجداول والمحتويات الأخرى. ومثل هذا الترميز عادة ما يتم تنفيذه عن طريق لغة تعرف بلغة التهيئة أو الترميز Markup language.

وقد انشغلت كثير من الجهات المهتمة بتمثيل النصوص لعدة سنوات بتطوير برنامج أو خطة لتهيئة النصوص أو ترميزها تعرف باللغة المعيارية الموحدة لتهيئة النصوص Standard Generalized markup Language (SGML). وتعد لغة تهيئة النص الفائق Hypertext markup Language (HTML) المستخدمة كصيغة عرض النصوص على الويب إحدى اللغات المشتقة من تلك اللغة المعيارية الموحدة.

وبما أن عرض إحدى الوثائق في صيغة اللغة المعيارية الموحدة أمر منفصل عن كيفية استخدام هذا النص، فإن هذا النص نفسه والذي تم تهيئته عن طريق علامات اللغة المعيارية الموحدة يمكن أن يقدم في أشكال وصيغ

(١) تعرف في الإنتاج الفكري بمسميات كثيرة منها لغات الترميز أو لغات التعليم أو التتويج (المترجمان) .

متعددة، كأن يعرض على ورق أو على أسطوانة مدمجة أو كنص مباشر أو كنص فائق إلى آخر هذه الأشكال. ولعل هذه الإمكانية هي التي جعلت هذه اللغة تلقى قبولاً لدى الناشرين الذين عادة ما يطمحون إلى إنتاج إصدارات متعددة من العمل نفسه، وقد كانت الإصدارات الجديدة من معجم أكسفورد الحديث للغة الإنجليزية من أولى تطبيقات استخدام اللغة المعيارية لتهيئة النصوص في هذا الصدد. كذلك حظيت هذه اللغة بالاستخدام المكثف من جانب الباحثين في مجال الإنسانيات الذين وجدوا فيها طريقة مناسبة لترميز بنية النصوص التي تكون مستقلة عن أي نظم آلية محددة أو حتى مجرد طريقة للعرض.

المكتبات الرقمية للدوريات العلمية :

التجارب المبكرة:

حظي التوجه نحو بناء مجموعات على الخط المباشر من الدوريات العلمية باهتمام العديد من الناشرين والمكتبات خلال نهاية ثمانينات القرن العشرين، وقد أزيلت تلك العقبات الفنية التي كانت تقف حاجزاً في طريق إنجاز تلك المشروعات منذ وقت مبكر، وإن كان أثرها لا يزال باقياً؛ فتكاليف الاختزان المباشر انخفضت، كما اتسع نطاق استخدام الحاسبات الشخصية والشبكات، بالإضافة إلى توافر برمجيات جيدة لإدارة قواعد البيانات، وقد تمثلت العقبات الكبرى في طريق بناء المكتبات الرقمية في وجود الإنتاج الفكري العلمي على الوسائط الورقية وليس في صيغ إلكترونية، كما كانت المؤسسات مهيأة للتعامل مع الوسائط المادية وليس مع شبكات الحاسبات الآلية.

ويعد مشروع عطارد أو ميركوري للمكتبة الإلكترونية Mercury electronic library project (الذي تبنته جامعة كارنيجي ميلون خلال الفترة من ١٩٨٧-١٩٩٣م) أول محاولات إنشاء مكتبة رقمية داخل إحدى المدن الجامعية. وقد استفاد هذا المشروع من البنية التحتية المتقدمة للحاسبات الآلية التي كانت تتمتع بها جامعة كارنيجي ميلون، والتي كانت تضم شبكة عالية القدرة، وقسمًا أكاديميًا ممتازًا لعلوم الحاسب، ناهيك عما يحسب لمكتبات تلك الجامعة من أعراف راسخة في رعاية الابتكارات. وبعد ذلك بقليل ظهر مشروع "كور Core" بجامعة كورنيل Cornell university الذي تركز اهتمامه على تجميع الصور من الدوريات الكيميائية، وقد دعمت جهود هذين المشروعين بمساعدة الناشرين العلميين بهدف إنشاء مجموعات من صور الصفحات المباشرة online page image وفي الوقت الذي تركز فيه اهتمام مشروع عطارد (ميركوري) على بناء نظام إنتاج، انصب اهتمام مشروع كور على البحث في واجهات المستفيدين والجوانب الأخرى للنظام من وجهة نظر الكيميائيين. وعلى الرغم من أن كلا المشروعين - ميركوري وكور - سعيًا إلى تحويل مقالات الدوريات الموجودة من الشكل المطبوع إلى الصور الإلكترونية أو ما تعرف بصور خرائط البتات أوالصور المرسومة خرائطياً Bit-mapped images، فإن هذا التحويل لم يتصور كمستقبل بعيد للمكتبات العلمية، وإنما يعكس- ببساطة- الوضع الذي لم يكن فيه بمقدور أي من ناشري الدوريات العلمية تقديم صيغ أخرى له. وإذا كانت المطابع قد استخدمت الحاسبات في محيط عملية الطباعة لعدة سنوات، فإن نظمها كانت مهيأة أساساً لإنتاج المواد المطبوعة، حيث كانت ملفات

الطباعة تتاح بتشكيلة متنوعة من أشكال الحروف، كما كانت عمليات مراجعة بروفات الطباعة تتم بشكل مستقل وليس على الملفات الأساسية ولذلك فإن هذه العمليات لا يمكن أن تستخدم في المكتبات الرقمية بدون جهود أخرى متعددة.

وقد تبعت مشروعي ميركوري وكور مشروعات أخرى كثيرة مستثمرة تقنية المسح الضوئي لمقالات الدوريات على هيئة صور، ومن أشهر هذه المشروعات مشروع توليب .Elsevier science publishing Tulip project الذي تبناه الناشر المعروف إلسفير، وعلى مدار ثلاث سنوات زود إلسفير مجموعة من الجامعات (بما فيها كارنيجي ميلون وكورنيل) بصور مأخوذة من ثلاث وأربعين دورية في مجال علم المواد Materials science. وقد قامت كل جامعة من هذه الجامعات بتحميل هذه الصور على حاسباتها،ثم إتاحتها على شبكتها المحلية.

ولم تكن مشروعات ميركوري وكور وتوليب بمثابة نظم إنتاج طويلة الأجل، فكل مشروع منها كان له مشكلاته الفنية، فضلاً عن المعاناة من قلة حجم المجموعة المقدمة للباحثين، غير أنها جسدت إمكانية تحقيق العوائد المحتملة للمكتبات الرقمية على المستوى الميداني.

وقد استطاعت الأجيال التالية من التطورات التي يشهدها قطاع النشر الإلكتروني أن تستفيد من إمكانات التخزين الآلي منخفضة التكاليف، والتي تكفل إمكانية الإتاحة المباشرة لمجموعات كبيرة. وأدى نشوء الويب ومتصفحاتها الواسعة الانتشار للذهاب بعيداً تجاه تبسيط التطوير في واجهات

المستفيدين. ومع أن متصفحات الويب ليست هي الأداة المثالية للتعامل مع المكتبات الرقمية، فإنها كانت بداية موفقة وكان لها - بالرغم من ذلك - فائدتها الكبرى في هذا الصدد؛ حيث كانت مناسبة لجميع الحاسبات ونظم التشغيل القياسية. ولم يعد حتماً أن يقوم كل مشروع بتطوير برنامج واجهة المستفيدين الخاصة به والخاصة بكل نوع من أنواع الحاسبات التي تستخدمها.

اللوحة رقم (٣ - ٢)

مشروعا ميركوري وكور **(Mercury and CORE)**

مشروع ميركوري من المشروعات الخمسية التي تهدف إلى إنشاء نموذج أولي لمكتبة رقمية في جامعة كارنيجي ميلون، وقد دشن هذا المشروع عام ١٩٨٨م، غير أن بدايته المباشرة كانت في عام ١٩٩١م. واستمر بعدها مخلفاً وراءه العديد من قواعد البيانات النصية وعددًا قليلاً من صور مقالات الدوريات في علوم الحاسب، وقد قدم هذا المشروع نموذجاً جيداً للوضع القائم قبل ظهور الويب.

وقد كان أحد أهداف المشروع نشر صور لمقالات الدوريات المرخص بها من قبل ناشريها، وقد تم تحديد أربعة ناشرين ممن ينشرون ست عشرة دورية من الدوريات العشرين المتخصصة في مجال الحاسب والتي تستخدم بكثافة بالجامعة. وكان هؤلاء الناشرون الأربعة هم: جمعية مصنعي الحاسبات ACM، ومعهد المهندسين الكهربائيين والإلكترونيين Institute of Electrical and Electronics Engineers، وإلسفير Elsevier، وبرجامون Pergamon.، مع ملاحظة أن إلسفير قام بشراء برجامون أثناء المشروع. ولم

يكن لأي من هؤلاء الناشرين إصدارات مقروءة آلياً من دورياتهم، لكنهم سمحوا بتحويل موادهم المطبوعة لأغراض الاستخدام داخل المكتبة. ومن ثم فقد كان الجزء الهام من العمل منصباً على عمليات التحويل والتخزين والتوزيع لصور الصفحات عبر شبكة المدينة الجامعية.

لقد كان النموذج الأصلي لمشروع ميركوري يتمثل في عملية بحث قواعد البيانات النصية بهدف تحديد المعلومات المراد عرضها، وقد اختيرت الإصدارة الأولى من معيار Z39.50 لتكون بروتوكول إرسال الاستفسارات بين الحاسبات العميلة والحاسبات الخادمة التي اختزنت بها الفهارس. وقد قدم مشروع ميركوري مفهوم الحاسب الخادم المرجعي "Reference Server" الذي يحتفظ بالمعلومات عن البيانات المختزنة في الحاسبات الخادمة الأخرى، والحقول التي يمكن بحثها، والكشافات، وقيود تداول المعلومات أو الوصول إليها. ومن أجل عرض الصور الخرائطية، قام مشروع ميركوري بتطوير خوارزمية جديدة لأخذ صور الصفحات المختزنة في صيغ مضغوطة، حيث يتم بموجب هذه الخوارزمية إرسال الصور عبر الشبكة، ثم إعادة فك ضغطها، ثم عرضها في وقت استجابة يتراوح ما بين ثانية وثانيتين للصفحة الواحدة.

ومنذ أن ارتبط مشروع ميركوري بالإنترنت، وقام الناشرون بالترخيص لمعظم المواد التي اختيرت، أصبح الأمن مطلباً ضرورياً؛ فالجامعة كان لديها بالفعل مجموعة ناضجة من الخدمات الشبكية والتي كانت تعرف بخدمات أندرو Andew. كما كان مشروع ميركوري قادراً على استخدام خدمات أندرو القياسية لأغراض التثبت من صحة المعلومات والطباعة،

وكانت المعلومات ترسل إلى الحاسبات الأخرى عبر البريد الإلكتروني.

أما مشروع كور CORE فقد كان مشروعًا مشتركًا بين بيل كور Belcore وجامعة كورنيل، وأو سي إل سي OCLC، والجمعية الأمريكية الكيميائية، وقد بدأ هذا المشروع خلال الفترة من ١٩٩١-١٩٩٥م، وتم بموجبه تحويل حوالي ٤٠٠,٠٠٠ صفحة تمثل المقالات المنشورة خلال أربع سنوات في الدوريات العشرين التي تنشرها الجمعية الكيميائية الأمريكية.

وقد أفاد مشروع كور من عدد من الأفكار التي باتت متداولة وشائعة في مشروعات التحويل، حيث كانت هناك نسختان من كل مقالة، واحدة مأخوذة عن طريق المسح الضوئي، والأخرى صورة نصية مهيأة عن طريق اللغة المعيارية الموحدة لتهيئة النصوص SGML. على أن تعمل الصورة المأخوذة عن طريق المسح الضوئي على ضمان ظهور الصفحة الأصلية الورقية بشكلها وتصميمها نفسيهما، عندما تعرض على الشاشة أو عند إعادة طباعتها، في حين تستخدم الصورة النصية المهيأة عن طريق اللغة المعيارية الموحدة SGML لبناء كشاف النص الكامل full-text index لأغراض استرجاع المعلومات ولأغراض العرض السريع على شاشات الحاسبات. كما كان يتم اختزان صورتين من الصور الممسوحة ضوئياً لكل صفحة، واحدة تستخدم لأغراض الطباعة والأخرى لأغراض العرض على الشاشة. بحيث تظهر الصورة المخصصة لأغراض الطباعة باللونين الأبيض والأسود بمعدل ٣٠٠ نقطة في البوصة، في حين كانت الصورة المخصصة للعرض تظهر بلون رمادي بمعدل ١٠٠ نقطة في البوصة.

وقد كان مشروع "كور" أحد المشروعات الأولى التي سعت لإثبات صحة كثير من النتائج التي انتهت إليها الدراسات المتخصصة، وكانت إحدى هذه النتائج أن المشكلات الفنية المتصلة بعمليات العرض والتخزين واختزان المواد العلمية المعقدة هي بالفعل مشكلات جوهرية، وخاصة إذا كانت هذه المواد مصممة في الأصل لكي تصدر مطبوعة، وجاءت عملية رقمنتها في مرحلة لاحقة. ومن ناحية ثانية، ألقى مشروع كور الضوء على مدى تأثير الحجم Scale، حيث شغلت مجموعات كور المكونة فقط من مقالات عشرين دورية ما يقرب من ثمانين جيجابايت من حيز التخزين، وقد كشفت جميع مشروعات الرقمنة المبكرة أن المشكلات التشغيلية operational problems لإدارة المجموعات الضخمة كانت أكبر مما كان متوقعاً لها.

وبرغم هذه العقبات، فقد أوضحت الدراسات التي أجريت على واجهة المستفيدين الخاصة بمشروع كور أن ثمة نجاحًا مرتقباً وواعداً ينتظر هذا المشروع. فالكيميائيون قد تآلفوا مع تلك المجموعات، مع أنهم قد يفضلون الإصدارات المطبوعة من المقالات حينما يكونون بصدد قراءتها كاملة بتفاصيلها. في حين تبين لهم أن الإصدارات المتاحة مباشرة online versions كانت أيسر بكثير في عمليات البحث، وأن ما يعرض على الشاشة كان يفوق احتياجاتهم. وأخيراً يعد مشروع كور من الدراسات الريادية التي أكدت أهمية عملية التصفح كبديل لعملية البحث، وهي تلك النتيجة التي أثبتتها الويب بشكل واضح فيما بعد.

الدوريات الإلكترونية :

ظل كبار ناشري الدوريات العلمية المعروفين مترددين حتى منتصف

التسعينات من القرن العشرين بشأن النشر على الخط المباشر online publishing، وبالرغم من أن النشر التجاري على الأقراص المدمجة قد قطع شوطاً كبيراً في هذا الاتجاه، إلا أن قليلاً من الدوريات هو الذي أتيح بشكل مباشر، فالناشرون لا يستطيعون تحقيق مكاسب مالية معقولة من وراء النشر المباشر - كما أنهم يتخوفون من الخسارة التي تلحق بهم من جراء إتاحة موادهم بهذه الطريقة، ثم بيعها مطبوعة بعد ذلك. وبحلول عام ١٩٩٥ م تقريباً كان من الواضح أن ثمة تغيرات كثيرة بدت في الأفق عن كيفية استخدام الناس للمعلومات. كما اتسعت مجالات إفادة كل من المكتبات والأفراد على حد سواء من خدمات المعلومات المباشرة، واستخدامهم للأشكال الأخرى من المعلومات الإلكترونية أكثر من استخدامهم للمصادر المطبوعة، كما كانت عملية نشر المواد المطبوعة عرضة للتنافس في ظل تناقص الميزانيات الثابتة للمكتبات.

ولأن معدلات توزيع الدوريات تظل في كساد أو انخفاض، فإن نشرها لا يحقق الأرباح المرجوة in a price spiral وقد استخدم الناشرون في الماضي التقنية الحديثة في خفض تكاليف الإنتاج، ومن ثم تخفيف حدة بعض الآثار السلبية للتوزيع المحدود، ولكن الآن بات تحقيق مثل تلك المدخرات أمراً منتهياً. ويدرك الناشرون أنه من الواجب عليهم حتى يتمكنوا من تنمية أوضاعهم المالية، أو على الأقل المحافظة على مستواها الحالي أن يمتلكوا منتجات تناسب السوق المتسعة للنشر المباشر حيث أصبح من المؤكد أن المعلومات الإلكترونية تبدو سوقاً واعدة ومتنامية.

لقد تحولت الغالبية العظمى من أكبر ناشري الدوريات العلمية بسرعة إلى

النشر الإلكتروني، وتبين اللوحة رقم (٣-٤) بالتفصيل التوجه الذي اتخذته جمعية مصنعي الحاسبات Association for Computing Machinery. وعلى هذا التوجه نفسه - مع تفاوت طفيف - يسير كل من إلسفير Elsevier وويلي wily والمطبعة الأكاديمية Academic press،هذا فضلاً عن بعض دور النشر التجارية الكبرى، والجمعية الكيميائية الأمريكية ACA، ومطبعة معهد ماساشوستيس للتكنولوجيا MIT press، ناهيك عن بعض مطابع الجامعات الأخرى.

إن مصطلح " دورية إلكترونية" يشيع استخدامه ليصف المطبوع الذي يحتفظ بالعديد من خصائص الدوريات المطبوعة، غير أن هذا المطبوع تتم إتاحته وتوزيعه على الخط المباشر. كما يستخدم أيضًا - وإن كان بصورة مربكة إلى حد ما- ليشير إلى كل من الدورية التي تنتج رقمية بشكل كامل، وتتاح على الخط المباشر فقط. كذلك يستخدم للإشارة إلى الإصدارة الرقمية من دورية تصدر في الأصل مطبوعة (مثل دورية The Association of computing machinery journal المشار إليها في اللوحة رقم (٣- ٤).

لقد قام عدد من الناشرين الضالعين في صناعة النشر بتقديم عدد قليل من الدوريات المتاحة كلية على الخط المباشر، وهناك جهود أخرى مشابهة في الاتجاه نفسه تبنتها جماعات أخرى. وقد أعدت بعض هذه الدوريات المتاحة على الخط المباشر لتحاكي مميزات وإجراءات إنتاج وتوزيع الدوريات التقليدية، وربما تعد " المجلة المباشرة للمحاولات الإكلينيكية الجارية The on-line journal of current clinical trails" التي أصدرتها الجمعية الأمريكية لتقدم العلوم بالتعاون مع أو سي إل سي OCLC - من

أكثر المجلات طموحاً في السير على هذا الاتجاه. وعلى عكس المطبوعات الإلكترونية الأخرى، فإن هذه الإصدارات المتاحة على الخط المباشر جاءت بمثابة إصدارات ثانوية تالية للإصدارات المطبوعة منها. وقد خطط لهذه الدورية الجديدة لتكون دورية محكمة بمعنى الكلمة، والتي ستكون الإصدارة النهائية منها هي الإصدارة الإلكترونية - مع ملاحظة أنه لم تكن هناك إصدارة مطبوعة-. وبما أن الناشر كان له حق التحكم الكامل في المجلة، فقد كان قادراً على تصميمها واختزانها في شكل يناسب عرضها وتداولها إلكترونياً. ومع ذلك لم تحظ هذه الدورية مطلقاً بقبول الباحثين أو الأطباء. ولقد ظهرت المجلة المباشرة للمحاولات الإكلينيكية الجارية أول مرة في عام ١٩٩٢م، غير أنها لم تستطع اجتذاب أعداد البحوث والدراسات الجيدة التي كان قد خطط لها أن تستوعبها، وهذا على أية حال هو مصير العديد من الأعمال الرائدة.

تحرص الدوريات الإلكترونية الأكثر حداثة على أن تجمع بين مميزات الدوريات التقليدية بالإضافة إلى الصيغ والخدمات التي توفرها إمكانات النشر الإلكتروني. وخير مثال على ذلك "مجلة المكتبات الرقمية D- Lib Magazine". التي أتحناها عام ١٩٩٥م على الخط المباشر لتغطي البحوث والإبداعات حول المكتبات الرقمية. ومن المعروف أن المنشورات المطبوعة مقيدة بمواصفات تصميم لا يمكن التنازل عنها، في حين أن المطبوعات المتاحة على الخط المباشر يمكن أن تسمح للمؤلفين أن يكونوا مبدعين في استخدامهم لتلك التقنيات. كما أن الدوريات التقليدية عادة ما تظهر في إصدارات تشمل كل واحدة منها مجموعة مقالات، هذا في الوقت الذي يمكن

فيه لبعض الدوريات الإلكترونية أن تنشر كل مقالة على حدة بمجرد أن تكون جاهزة للنشر، مع أن "مجلة المكتبات الرقمية" هذه اعتادت أن تنشر إصدارات شهرية مع الحرص الشديد على الصدور في موعدها وبدون تأخير. وأنه لمؤشر على تعاقب الأيام وتغيرها أن تنشأ هذه المجلة التي تتاح بكاملها على الخط المباشر وبدون اعتمادها على نظام التحكيم كدورية هامة متخصصة في مجال المكتبات الرقمية.

اللوحة رقم (٣ - ٣)

مشروع مطبعة هاي وير **(Highwire Press)**

هي أحد المشاريع التي جازفت بها مكتبات جامعة ستانفورد، حيث وضعت أفضل الدوريات العلمية والطبية على الخط المباشر من خلال عمل شراكة بين الهيئات العلمية والمهنية التي تُصدرها. ويعزا نجاحها هذا إلى مدى الاهتمام الكبير برغبات أكبر الباحثين المتخصصين وبتركيزها على أهم الدوريات وامتيازها الفني.

بدأت مطبعة هاي وير نشاطها في عام ١٩٩٥م بإصدار نسخة على الخط المباشر لدورية الكيمياء الحيوية، والتي يضم مجلس تحريرها أعضاء من هيئة التدريس بجامعة ستانفورد. ويحتوي كل عدد أسبوعي من هذه الدورية على ٨٠٠ صفحة، حيث لا يوجد من يقوم بقراءتها من الغلاف إلى الغلاف، ولذلك فإن واجهة التعامل لمطبعة هاي وير تتعامل مع المقالات على اعتبار أنها مستندات منفصلة. كما أن هناك طريقتين للوصول للمقالات، إحداهما تصفح المحتويات من خلال الأعداد الأسبوعية والثانية

من خلال البحث. وخيارات البحث هذه تشمل البحث عن المؤلفين، والبحث عن كلمة في موضوع المقال، والبحث عن ملخص محتوى النص، والبحث عن نصوص جميع المقالات. كما صممت شاشة البحث وطريقة عرض المقالات بأسلوب يؤكد على أنها دورية الكيمياء الحيوية وليست دورية منشورة بواسطة مطبعة هاي وير؛ حيث تبذل جهود عظيمة لعرض الحروف الإغريقية، والرموز الرياضية والحروف الخاصة الأخرى.

وفي خلال ثلاث سنوات، تخطت مطبعة هاي وير مرحلة التجربة إلى المرحلة العملية بمئات الدوريات على الخط المباشر بما في ذلك الدوريات المتخصصة في العلوم. وبالرغم من مواردها فقد خاضت الجمعية الأمريكية لتطوير العلوم The American Association for the Advancement of Science (AAAS) - الناشر لدورية العلوم - تجربة الخوف التي تواجهها أي جمعية تصدر عددًا قليلاً من الدوريات، ويتمثل الخوف عندما تم وضع دورية العلوم على الخط المباشر بأنه ربما يمتد ذلك إلى الموظفين مما قد يؤثر على الجودة العامة لعملهم. وقد ساعدت الشراكة بين مطبعة هاي وير والجمعية الأمريكية لتطوير العلوم على مشاطرة تكاليف التطوير مع دوريات هيئات أخرى، وكذلك التعاون مع المتخصصين في جامعة ستانفورد. وقد أصبحت الجمعية الأمريكية راضية عن عدد الأشخاص الذين يزرون موقعها أسبوعياً، حيث إن معظمهم ليسوا من القراء المنتظمين لدورية العلوم.

اللوحة رقم (٣ - ٤)

المكتبة الرقمية لجمعية مصنعي الحاسبات

(The Association for Computing Machinery's Digital Library)

هي جمعية مهنية تنشر سبع عشرة دورية متخصصة في مجال علوم

الحاسب الآلي. بالإضافة إلى ثمان وثلاثين مجموعة ذات اهتمامات متخصصة تجري العديد من المؤتمرات والتي تُنشر معظم أبحاثها. كما أن أعضاءها هم من الخبراء في مجال علوم الحاسب ومن بينهم العديد من الأشخاص الذين ساهموا في بناء شبكة الإنترنت والويب. ويعتبر هؤلاء الأشخاص من الأوائل الذين اعتادوا على التواصل على الخط المباشر، والذين توقعوا بأن تصبح جمعيتهم هي الرائدة للتحول إلى الدوريات المباشرة على الإنترنت.

وكما هو معتاد فقد صدرت النسخة التعريفية من الدورية في شكل مطبوع. أما في عام ١٩٩٣م، فقد قررت جمعية مصنعي الحاسبات ACM أن تجري عمليات إنتاجها المستقبلية عن طريق استخدام نظام حاسب آلي يقوم بإنشاء قاعدة بيانات لمقالات الدورية، وجلسات عمل المؤتمرات، والمجلات والصحف بحيث تكون جميعها على صيغة اللغة المعيارية الموحدة لتهيئة النصوص SGML. ومؤخراً قررت جمعية مصنعي الحاسبات ACM كذلك تحويل عدد كبير من دورياتها القديمة لبناء مكتبة رقمية تحتوي على منشورتها من عام ١٩٨٥م وما بعده.

وقد استخدمت ملفات اللغة المعيارية الموحدة لتهيئة النصوص SGML لإنتاج الدوريات المطبوعة، ومع ذلك فقد خططت جمعية مصنعي الحاسبات لليوم الذي سيحتاج فيه الأعضاء لاسترجاع المقالات مباشرة من قاعدة البيانات الموجودة على الخط المباشر وقراءاتها على شاشة الحاسب الآلي أحياناً، وتحميلها على طابعة محلية أحياناً أخرى. ويمكن للمكتبة الحصول على ترخيص لأجزاء من قاعدة البيانات تلك، أو أن يكون لديها اشتراك عام

لموظفيها.

وتستخدم المجموعات التي دشنت على الخط المباشر في عام ١٩٩٧م واجهة الويب التي تقدم للقراء فرصة تصفح محتويات صفحات دوريات جمعية مصنعي الحاسبات، بالإضافة إلى إمكانية البحث عن طريق المؤلف والكلمات المفتاحية. ثم يستطيع المشترك قراءة النص كاملاً عند تحديد المقالة، في حين يتطلب من القراء الآخرين دفع رسوم للوصول للنصوص الكاملة، ولكن قد يكون بإمكانهم قراءة ملخصات النصوص بدون دفع أية رسوم.

وتحتاج جمعية مصنعي الحاسبات العائد المالي لتغطية تكاليف المكتبة الباهظة ولكنها في الوقت نفسه لا ترغب في تقييد المؤلفين والقراء بصورة مفرطة. وتقع ترتيبات العمل التجاري المبدئي في فئتين: إحداهما مرتبطة بالمؤلفين، والأخرى بالقراء.

وقامت جمعية مصنعي الحاسبات بنشر سياسة حقوق الطبع والنشر في عام ١٩٩٤م، والتي تصف العلاقة بين الجمعية (كناشر) والمؤلف. كما حاولت الموازنة بين حقوق المؤلفين واحتياجات الجمعية لخلق عائد مالي من منشوراتها. وفي ذلك الوقت نشرت جمعية مصنعي الحاسبات سياستها الجديدة على خادم الويب الخاص بها. ومن المميزات الهامة في هذه السياسة هي الإقرار الواضح والصريح بتوزيع معظم مقالات الدورية على الويب أولاً.

وفرضت جمعية مصنعي الحاسبات رسوماً للوصول للنص الكامل للمقالة بهدف تحقيق عائد مالي؛ إذ قد يقوم الأعضاء والمكتبات بالاشتراك في

الإصدارات الإلكترونية لدوريات الجمعية والتي كانت أسعارها أقل من أسعار الإصدارات المطبوعة بحوالي ٢٠%، وكحل بديل فإنه بإمكان الأفراد الدفع لمقالات فردية محددة. وتهدف بنية الأسعار إلى تشجيع المشتركين للاشتراك في مجموعة كاملة من المطبوعات بدلاً من قراءة دوريات مستقلة فقط.

مشروعات التحويل والإنسانيات :

يعد حفظ مواد اليوم كذاكرة بعيدة الأمد للغد إحدى المهام الأساسية التي تضطلع بها المكتبات البحثية. فالمكتبات الكبرى تمتلك مجموعات قيمة تشكل المادة الخام لدراسة التاريخ والإنسانيات، وهذه المواد عادة ما تتكون في الأساس من المواد المطبوعة وغيرها من المصنوعات المادية الأخرى physical artifacts. وقد أدى تطور المكتبات الرقمية إلى إثارة حماس كبير نحو رقمنة هذه المجموعات، وكان أحد دوافع هذا الحماس أن المواد القديمة غالباً ما تكون في حالة مادية هشة، وأن عمل نسخ رقمية من هذه المواد يؤدي إلى حفظ محتواها، فضلاً عن إتاحتها للجميع.

ويسعى عدد من مشروعات المكتبات الرقمية إلى تحويل الوثائق الورقية المتوافرة لها إلى صيغة الصور الإلكترونية أو ما تعرف بصور خرائط البتات أو الصور المرسومة خرائطياً Bit-mapped images. ويجري الآن مسح الوثائق المطبوعة صفحة صفحة، في حين كانت الصفحة في التجارب المبكرة تسجل كسلسلة من النقاط البيضاء والسوداء (غالباً ما تكون ٣٠٠ نقطة في البوصة). وحديثاً بلغ المسح الضوئي الملون عالي الوضوح أعلى درجات الدقة. وعادة ما تكون صور خرائط البتات الجيدة واضحة بما فيه الكفاية

لدرجة أنه يمكن عرضها على شاشات الحاسبات الكبيرة، أو طبعها على الورق في صورة مقروءة جيدة good legibility. ولأن عملية المسح الضوئي عالي النقاء تنتج عدداً كبيراً من النقاط في الصفحة الواحدة، فهناك العديد من الطرق المختلفة المستخدمة لضغط الصور بهدف التقليل من عدد البتات التي ستختزن ومن حجم الملفات التي سترسل عبر الشبكة. ولكن حتى أبسط الصور مكونة من عشرات الآلاف من البتات لكل صفحة.

ونظراً لأن المجالات العلمية الفنية كانت هي القوة المحركة في مجال المكتبات الإلكترونية وخدمات المعلومات، فإن الاحتياجات الأساسية للمجالات العلمية الأخرى مثل مجموعات حروف اللغات غير الإنجليزية قد أهملت في الغالب، وباتت الإنسانيات عرضة للتجاهل والنسيان. غير أن الجيل الجديد من الباحثين في مجال الإنسانيات يتابعون عن كثب تطبيقات استخدام الحاسبات، ومن حسن حظهم أن لديهم أصدقاء في هذا المجال. وفي اللوحة رقم (٣-٦) وصف لمشروع جستور JSTOR [١]، الذي تبنته مؤسسة أندرو دبليو ميلون Andrew W. Mellon الذي يسعى إلى توفير نفقات المكتبات والحصول على الإنتاج الفكري الهام من الدوريات وتقديمها لأكبر عدد من المستفيدين. وبشكل عام لا يمكن لهذا المشروع أن يتحقق إلا بوجود المكتبات الرقمية.

المكتبات والابتكار :

من المعروف أن المكتبات بطبيعتها منظمات محافظة؛ فمجموعاتها

(١) جستور JSTOR هي اختصار لجملة اختزان الدوريات Journal Storage (المترجمان) .

وفهارسها تطورت على مدار عقود أو قل على مدار قرون، كما أن الخدمات المستحدثة تقدم فيها على حذر لأنه يتوقع لهذه الخدمات أن تبقى لفترة طويلة. ومع ذلك فإن المكتبات في المجالات الفنية كانت دائماً مغامرة، فمارك MARC ، وأو سي إل سي OCLC، ومشروع النظم المترابطة، وميركوري Mercury، وكور CORE، ومشروعات التحويل الحديثة، لم تخترع كل تلك التقنيات التي تستخدمها، ولكن الانتشار واسع النطاق والطابع العملي لمثل تلك النظم هو الذي جعلها رائدة في المجال.

وقد ناقشنا في الفصل الثاني المجتمع الذي نما وترعرع حول الإنترنت والويب، غير أن كثيراً من أعضاء هذا المجتمع اكتشفوا المعلومات المتاحة على الخط المباشر متأخرين جداً، كما أنهم لم يدركوا أن المكتبات ودور النشر قد طورت كثيراً من المفاهيم التي أصبحت ركائز للمكتبات الرقمية قبل ظهور الويب بسنين.

اللوحة رقم (٣ - ٥)

الذاكرة الأمريكية وبرنامج المكتبة الرقمية الوطنية

(American Memory and the National Digital Library Program)

تمتلك مكتبة الكونجرس - أكبر مكتبات العالم - مجموعات خاصة قيمة من المواد الفريدة وغير المنشورة، ومن بين تلك الكنوز الأوراق الشخصية الخاصة بثلاثة وعشرين من الرؤساء الأمريكيين، كذلك يتوافر بها العديد من الكتب النادرة والنشرات والدراسات التي تعد من المواد الهامة لدراسة الأحداث والحقب والحركات التاريخية. كما تعكس الملايين من الصور

الفوتوغـرافية والمطـبوعات والخـرائط والـنوت الموسـيقية والتـسجيلات الصوتية والصور المتحركة بجميع أشكالها توجهات الدولة،كما تمثل الناس والأمـاكن. وحتـى وقت قريب لم يكن أمام من يرغب في استخدام تلك المواد والإفادة منها سوى زيارة تلك المكتبة في مبناها الكائـن بكابيتل هيل Capital Hill.

وبـرنامج الذاكـرة الأمـريكية مـشروع ريـادي، بـدأ فـي عـام ١٩٨٩م واستمر حتـى عـام ١٩٩٤م بهدف إعـادة تكـوين مجموعات مختارة من أجل تـوزيعها أو بثها على المستوى الوطني في صيغ آلية أو محسبة. علـى أن يتم اختيار تلك المجموعات اعتماداً علـى مدى قيمتها لدراسة الثقافة والتاريخ الأمريكـي، واختيارها كذلك بهدف استكشاف مشكلات التعامل مـع المواد مـتعددة الأشـكال كالمـواد المطـبوعة والمـصورة والـصور المتحـركة والتـسجيلات الـصوتية والوثائـق النـصية. وقـد اسـتخدم بـرنامج الذاكـرة الأمـريكية فـي أول الأمـر تولـيفة مـن تمثـيلات مـرقمنة digitized representation محملـة علـى أقـراص مدمجـة، وصـيغ تناظـرية Analog forms علـى أشـرطة الفيديو، وإن كان قد أتيحت في أبريل من عام ١٩٩٤م ثلاث مجموعات من الصور الفوتوغرافية على الويب.

وعلـى أسـاس النجاح الذي حققه برنامج الذاكرة الأمريكية، بني برنامج المكتبة الوطنية الرقمية الذي تمثلت أهدافه في تحويل ملايين العناوين إلى مواد رقمية وإتاحتها عبر الإنترنت. وتركـز اهتمام البرنامج في أول الأمر على المواد الهامة المتصلة بتاريخ الولايات المتحدة الأمريكية، مثل مذكرات

والت ويتمان Walt Witman ووثائق المؤتمر القاري والاتفاقيات الدستورية the continental congress and the constitutional conventions.

وتعد بعض المجموعات التي تم تحويلها بمثابة أرشيف مترابط مثل الأوراق والصور الخاصة بشخصية واحدة أو بمنظمة محددة، وبعضها الآخر عبارة عن مجموعة من العناوين الموجودة في أشكالها الأصلية مثل المواد المطبوعة على النحاس daguerreotypes [١]، والصور الموجبة الورقية للأفلام المبكرة،كما أن البعض الآخر عبارة عن توليفات موضوعية قام بإعدادها بعض أمناء المكتبات والمتاحف أو بعض العلماء، ويتم اختيارها إما من داخل المجموعات الأرشيفية أو من مجموعات المكتبة كلها.

وقد اكتشف برنامج الذاكرة الأمريكية أن المعلمين بالمدارس كانوا متحمسين للوصول إلى هذه المواد المصدرية الأصلية، كما يرى برنامج المكتبة الوطنية الرقمية أن التعليم هو رسالته الأساسية. ويعد بحث الكشافات من أشمل وسائل الوصول إلى تلك المجموعات. ومع ذلك فلا يوجد للعديد من المجموعات الأرشيفية التي تم تحويلها سجلات للفهارس تعرف بكل عنوان فيها، وإنما تتوافر لها ما تعرف بمعينات البحث Finding aids، وهي عبارة عن وثائق منظمة تصف المجموعة كلها وصفاً شاملاً بما تشتمل عليه من عناوين دون الوصف التفصيلي لكل عنوان على حدة. ومن ثم فإن

---

(١) يعرف هذه الأسلوب بالتصوير الداجيري نسبة إلى مخترعه الفرنسي لويس داجيير، وهو أول طريقة ناجحة لطبع صور على لوح من النحاس المحسس بأملاح الفضة (المترجمان: عن حسب الله، ص ٧٣٦) .

الوصول إلى أي عنوان من العناوين التي اشتمل عليها مشروع الذاكرة الأمريكية عبارة عن توليفة من الطرق تشمل بحث السجلات الببليوجرافية بحثاً عن عناوين محددة تتوافر لها هذه السجلات، وتصفح المصطلحات الموضوعية، وبحث النصوص الكاملة، ثم بحث معينات البحث في المستقبل.

ويعد هذا البرنامج من البرامج الهامة من الناحية الفنية بسبب معدلات النجاح التي حققها وبسبب رؤيته الواضحة للأمور، وإدراكاً منها لطبيعة المشكلات طويلة الأمد المرتبطة بحفظ المجموعات الضخمة، ركزت مكتبة الكونجرس تركيزاً كبيراً على كيفية تنظيم عناوين مجموعاتها، وقد أخذ برنامج المكتبة الرقمية الوطنية عملية تحويل تلك المواد القديمة إلى صيغ رقمية بشكل جاد، واختيار الصيغة الأكثر ملاءمة لتمثيل المحتوى وممارسة ضبط الجودة الدقيق. ومن الملاحظ أن المواد النصية عادة ما يتم تحويلها مرتين، مرة كصورة مأخوذة عن طريق المسح الضوئي، ومرة أخرى للنص عن طريق تهيئته باللغة المعيارية الموحدة لتهيئة النصوص (SGML)، وقد تم عمل العديد من الصور من كل صورة فوتوغرافية، بحيث تتراوح ما بين الصورة الصغيرة المنمنمة منخفضة النقاء thumbnail إلى الصورة عالية النقاء لأغراض الحفظ الأرشيفي.

ومن الملاحظ كذلك أن عدداً من المواد التي تم اختيارها لأغراض التحويل كانت مواد لا تتمتع بحقوق الملكية الفكرية، وليس على توزيعها أية قيود أخرى، في حين كانت ثمة قيود على بعض المواد الأخرى. وفضلاً عن حقوق التأليف، كانت هناك حالات طلب الأذن بتحويلها من مالكي أصولها،

وفـي حالة المواد الأقدم - وخاصة المواد غير المنشورة - غالباً ما يكون من غير الممكن اكتشاف كل القيود التي يمكن تطبيقها بشكل ممكن تقبله، أو أنها مرتفعة الثمن بشكل يصعب معه إجراء بحث مضن عن كل عنوان فردي من عناوينها. ولذلك يجب على الموظفين القانونيين بالمكتبة أن يطوروا سياسات وإجراءات توازن بين قيمة إتاحة المواد في مقابل المخاطرة بانتهاك بعض الحقوق عن غير قصد.

إن بعض المكتبات تنظر إلى مكتبة الكونجرس باعتبارها صاحبة القيادة، حيث تعمل كعضو نشط في العديد من التكتلات المتخصصة، كما أنها أشرفت على برنامج منح هام، وأصبحت هي الراعية لبحوث المكتبة الرقمية، وإن الخبرة التي اكتسبتها المكتبة بالإضافة إلى مشاركتها الفعالة في العديد من المجموعات الفنية القيادية مكنتها من مساعدة مجتمع المكتبات بأسره على التقدم.

اللوحة رقم (٣ - ٦)

مشروع جستور **(JSTOR)**

جستور هو مشروع تبنته مؤسسة أندرو دبليو ميلون بهدف تزويد المكتبات الأكاديمية بالأعداد القديمة من الدوريات الهامة، وهو بذلك يجمع ما بين الأهداف الأكاديمية والأهداف الاقتصادية؛ فمن الناحية الأكاديمية يهدف هذا المشروع إلى بناء أرشيف يعتمد عليه للدوريات العلمية الهامة وتوفير مقومات الوصول إلى هذه الدوريات على نطاق واسع،أما من الناحية الاقتصادية يسعى المشروع إلى توفير نفقات المكتبات عن طريق إعفاء كل

مكتبة من حفظ المواد نفسها وصيانتها التي تتولى حفظها المكتبات الأخرى.

وقد رتبت مجموعات مشروع جستور موضوعياً حسب تخصصاتها - الاقتصاد، والتاريخ، والفلسفة - وهكذا، وقد قدر للمرحلة الأولى من المشروع أن تضم ما يقرب من مائة دورية في حوالي خمسة عشر مجالاً موضوعياً. وقد تكونت المجموعة بالنسبة لكل دورية من جميع الإصدارات من أول عدد حتى العدد الذي كان قد صدر قبل خمس سنوات من التاريخ المحدد للمشروع.

وقد تأسس مشروع جستور في أغسطس من عام ١٩٩٥م كمؤسسة مستقلة غير ربحية، تهدف إلى أن تصبح معتمدة على نفسها في التمويل من خلال الرسوم التي تفرض على المكتبات التي ترغب في استخدام قاعدة بيانات الدوريات، علماً بأن الرسوم التي تفرض ستكون أقل من تكاليف عمليات اختزان النسخ الورقية من هذه الدوريات.

وكان لهذه المؤسسة ثلاثة مكاتب: الأول في نيويورك ويتولى الأنشطة الإدارية والقانونية والمالية من حيث إدارتها والتنسيق بينها وعلاقتها بالمكتبات والناشرين، أما المكتبان الآخران فأحدهما في جامعة ميتشجان، والآخر في جامعة برينستون، ويضطلع كل واحد منهما بحفظ نسخة من قاعدة البيانات، هذا فضلاً عن رعاية البنية الفنية للمشروع وتطويرها، وتقديم الدعم الفني للمستفيدين، ثم الإشراف على عمليات التحويل من النسخ الورقية إلى الصيغ الإلكترونية، مع ملاحظة أن عمليات المسح الضوئي والحفظ يقوم بها موردون خارجيون. وحديثاً دشنت جستور نسخة ثالثة من قاعدة البيانات في جامعة مانشستر في المملكة المتحدة.

ويملك جستور حقوق ترخيص سارية المفعول مع الناشرين والمؤسسات المشتركة الأخرى. وفي سبيل تأكيده على الإصدارات القديمة يحرص مشروع جستور على عدم الدخول في منافسة مع الناشرين الذين تعتمد مكاسبهم المالية أساساً على الإصدارات الجديدة. وإلى الآن يعمل المشروع على تأمين الوصول للمجموعات لصالح المكتبات الأكاديمية التي تشترك في المجموعة الكاملة فقط، بحيث يعتمد في تقدير الرسوم التي تدفعها كل مكتبة على حجم هذه المكتبة. وتسهيلاً على هذه الجهات تتاح قوائم الرسوم ونماذج التراخيص مجاناً وبشكل مباشر على الإنترنت.

وفي سياق هذا المشروع يتم المسح الضوئي لجميع صفحات إصدارات الدوريات التي يتم الاحتفاظ بها وبدرجة عالية من الوضوح (٦٠٠ بت في البوصة) مع الحرص على ضبط الجودة. وعلى عكس بعض المشروعات الأخرى يتم اختزان نسخة واحدة فقط من كل صورة، أما النسخ الأخرى كالنسخ المنمنمة قليلة الوضوح Low-resolution thumbnails فلا تختزن، غير أنه يمكن التعامل معها آلياً عندما تُطلب. كذلك يتم تحويل النصوص عن طريق عملية " التعرف إلى الحروف بصرياً Optical Character Recognition" مع إخضاع النص لعملية تصحيح مكثفة، على أن يستخدم النص الذي تم تحويله لأغراض التكشيف فقط، وبالإضافة إلى ذلك يتم إنشاء ملف من قائمة المحتويات لكل مقالة. ويشتمل هذا الملف كذلك على البيانات الببليوجرافية مع الكلمات المفتاحية Keywords والمستخلصات إذا كانت موجودة.

الفصل الرابع

الابتكار والبحث

الابتكار موضوع متداول في عالمي المكتبات الرقمية وصناعة النشر الإلكتروني، وتأتي بعض الابتكارات تلبية لحاجة السوق إلى المعلومات، في حين يأتي البعض الآخر نتيجة لجهود البحث العلمي الذي تقوم به الجامعات أو تتبناه الشركات. ويسعى هذا الفصل إلى إلقاء الضوء على نشاط البحث العلمي، كما يقدم صورة مجملة لواقع البحوث الجارية في هذا المجال. أما التفاصيل الكاملة عن معظم القضايا التي يعرضها هذا الفصل فسوف نتناولها في فصول لاحقة.

كانت الابتكارات التي قدمتها المكتبات والناشرون حتى وقت قريب أبعد ما تكون عن المنهجية، يضاف إلى ذلك أن الغالبية العظمى منها كانت تفتقر إلى الميزانيات التي تدعم جهود البحث والتطوير. ويستثنى من ذلك كل من المكتبة القومية للطب The National Library of Medicine، ومركز الحاسب الآلي للمكتبات على الخط المباشر OCLC، إذ يقومان بإجراء البحوث بالمعنى الصحيح للكلمة من أجل استكشاف الحقائق دون تصورات مسبقة عن كيفية نشر تلك الأفكار في صورتها النهائية مستقبلاً. أما المكتبات الأخرى وغيرها من دور النشر فلها أساليبها في الابتكارات، والتي تمليها الاعتبارات العملية عن كيفية تعزيز خدماتها ومنتجاتها على المدى القريب. ولاشك أن معدل الابتكار عادة يتحكم فيه الوضع التقني ومدى توافر التمويل، بالإضافة إلى فهم الطلب الحالي على الخدمات وترجمته.

ومن الملاحظ أن الناشرين لم يقدموا بحوثاً كثيرة في هذا المجال، مع أنهم مؤهلون تأهيلاً جيداً يكفل لهم القدرة على طرح أفكار جديدة على السوق. وبالرغم من توافر الموارد لكبار الناشرين وبالدرجة التي تمكنهم من مباشرة الاستثمارات الضخمة، إلا أنهم يرون في المشروعات المتقدمة تطويرًا لمجال الأعمال وليس كبحوث علمية. وإن التحول نحو إتاحة الدوريات العلمية على الخط المباشر لا يمكن أن يسمى بحثاً علمياً، هذا بالرغم من أن هذه العملية قد غيرت مجرى صناعة النشر برمتها، لأن الدوريات التي يمكن أن تنتج في شكل مطبوع وعلى الخط المباشر ليست بالمهام اليسيرة بالتأكيد.

وتميل المكتبات لأن تكون أكثر توجهاً نحو الابتكار من الناشرين، مع أنها غالباً ما تبدو أقل تنظيماً في توجهها نحو التغيير بشكل عام. فهي عادة ما تنفق جل ميزانيتها على الأنشطة اليومية التي تمارسها، وغالباً ما ينظر إلى الابتكار فيها على أنه من قبيل الترف، وليس على أنه مفتاح إلى عالم المستقبل [١]. ومع أن المكتبات الكبرى تمتلك ميزانيات ضخمة، فإن نظم إدارتها لتلك الميزانيات لا تزال تفتقر حتى الآن إلى المرونة حتى أن أنشطة البحث والابتكار فيها تفتقر افتقاراً كبيراً إلى توفير الموظفين المؤهلين لهذا النشاط.

وربما يعد برنامج المكتبة الوطنية الرقمية الذي تبنته مكتبة الكونجرس من أهم المشروعات المكتبية في الولايات المتحدة التي تعكس مدى توسع المكتبات لتخرج عن حدود دورها التقليدي من خلال إدارة مجموعاتها بطريقة

(١) قد نختلف مع المؤلف في هذا الصدد (المترجمان).

مختلفة، ومن خلال التعامل مع أشكال جديدة من المواد، وإتاحة مجموعاتها على نطاق واسع، ولأن مكتبة الكونجرس لم تقدم سوى الدعم المالي المحدود لهذا البرنامج، فإنه يسعى إلى جمع تمويله من المؤسسات الخاصة، وعن طريق الهدايا والهبات، وعادة ما يكون التوظيف بهذا البرنامج بنظام عقود العمل قصيرة الأجل. وعلى الرغم من أن مكتبة الكونجرس دعمت بالفعل التطورات الفنية مثل دعمها لبروتوكول Z39.50، ومارك MARC، فإن مشروعها هذا الأكثر مثالية لا يحظى بالدعم المالي طويل الأمد.

وبرغم هذا التحفظ الواضح فإن المكتبات - وخاصة الجامعية منها- تتجه إلى الأمام صوب التوسع. وبالرغم من أن الهياكل التنظيمية للمكتبات الجامعية عادة ما تسعى إلى تثبيط التوجيهات نحو التغير الجذري، فإن أمناء تلك المكتبات الجامعية يعرفون أن استخدام الحاسبات هو أمر أساسي لمستقبل الاتصال العلمي، وينظر في هذا السياق إلى علماء الحاسبات وكأنهم أشبه بالأرانب البرية التي تركض في حلبة المكتبات الرقمية؛ إذ يتواثبون دفعة إلى الأمام ليكتشفوا مجالات جديدة، ثم يقفزون ثانية إلى مجالات أخرى. أما المكتبات الكبرى فهي أشبه بالسلاحف؛ تتحرك ببطء شديد ولكن كل خطوة تخطوها هي تمهيد للخطوة التالية. وغالباً ما تأخذ هذه الخطوات شكل المشروعات البؤرية Focused project، فتكون أحياناً على هيئة شراكة مع منظمات أخرى، وغالباً ما تكون ممولة من قبل مؤسسات، وقد لا يكون لدى المشروعات الفردية الرؤية نحو مبادرات البحوث الممولة تمويلاً حكومياً ضخماً، غير أن أثرها التراكمي ربما يكون كبيراً على المدى البعيد.

لقد نشأت العديد من المشروعات الهامة في كنف المكتبات الجامعية،

وتقوم بعض المكتبات الجامعية بتحويل المواد إلى الصيغ الرقمية؛ في حين يعمل بعضها الآخر مع الناشرين لإتاحة موادها على الخط المباشر. وقد قامت مطبعة هاي واير High Wire في جامعة ستانفورد بإتاحة الدوريات العلمية على الخط المباشر، كما قامت بعض المكتبات الجامعية بالتعاون مع إلسفير Elsevier Science في مشروع توليب TULIP بهدف استكشاف إصدارة مرقمنة للدوريات العلمية. وكما تتعاون كل من جامعتي ميتشجان وبرينستون في مشروع جستور JSTOR لتحويل الأعداد القديمة من الدوريات العلمية الهامة. وتعاونت جامعتا رتجرز وبرينستون في إنشاء مركز النصوص الإلكترونية في الإنسانيات The Center of Electronic Texts in the Humanities.

إن الغالبية العظمى من المشروعات تنشأ من خلال المنح التي تقدمها المؤسسات أو المصانع أو الحكومة الفيدرالية، فعلى سبيل المثال قام مشروع ميركوري في جامعة كارنيجي ميلون على المنح التي قدمها اتحاد الكنيسة الخيرية The Pew charitable trust، وشركة التجهيزات الرقمية، وهيئة الدفاع للمشروعات البحثية المتقدمة المعروفة بداربا The defense Advanced Research Project Agency (DARPA)، هذا فضلاً عن الدعم القليل المرحب به من قبل متبرعين آخرين. وهناك العديد من المؤسسات الخاصة التي تعد من المؤسسات الداعمة القوية لأنشطة المكتبات الرقمية في مجال الإنسانيات، من أهمها مؤسسة أندرو دبليو ميلون Andrew W.mellon، واتحاد جي باول جريتي J.Paul Gretty. وأخيراً وبالرغم من ميزانيتها الصغيرة إذا ما قورنت بميزانية كل من المؤسسة الوطنية للعلوم National Science Foundation وهيئة

الدفاع للمشروعات البحثية المتقدمة (داربا)، استطاعت هيئة الوقف الوطني للإنسانيات The National Endowment for The Humanities، أن تخصص نسبة معقولة من إعتماداتها المالية لحساب المكتبات الرقمية.

إن علماء الحاسبات يأخذون البحث العلمي مأخذ الجد، حيث اعتادوا على المشروعات طويلة الأجل، والتي لا تكون نتائجها النهائية في شكل منتجات أو خدمات، وإنما بمثابة مفاهيم جديدة، أو فهم أعمق للمجال. ومن المعروف أن معظم الأموال الخاصة بدعم البحوث في الولايات المتحدة تأتي من الهيئات الفيدرالية، وتعد أكبر هيئة دولية داعمة لنشاط البحث في مجال علوم الحاسب هي هيئة الدفاع للمشروعات البحثية المتقدمة (التي غيرت اسمها من أربا ARPA إلى داربا DARPA، ثم عادت مرة أخرى إلى أربا، ثم عادت مرة ثالثة إلى داربا). تأتي بعدها المؤسسة الوطنية للعلوم (NSF). وتعد داربا (DARPA) فرعاً من وزارة الدفاع الأمريكية، وعلى الرغم من أن مهمتها تنصب في نهاية الأمر في خدمة الأنشطة العسكرية ودعمها، فإنها عادة تتخذ نظرة شمولية، حيث تقوم بتشجيع البحوث الأساسية في جميع مجالات الحاسب الآلي، وتركز بشكل خاص على المشروعات البحثية التي تبني نظماً تجريبية كبرى. أما المؤسسة الوطنية للعلوم (NSF) فهي التي تضطلع بالمسؤولية الأولى في تطوير العلوم والهندسة في الولايات المتحدة الأمريكية. وتستنفذ برامجها سنوياً أكثر من ٣,٣ بليون دولار أمريكي لدعم ما يقرب من ٢٠,٠٠٠ مشروع بحثي وتعليمي، وهي تدعم مشروعات البحوث في مجال علوم الحاسب الآلي وتطبيقاته في جميع التخصصات العلمية تقريباً.

وتتوافر للعديد من الشركات الهندسية وشركات الحاسب الآلي ميزانيات

ضخمة مخصصة لبحوث استخدام الحاسبات وتطويرها. ويخصص ما يقرب من ١٠% من إجمالي الميزانية لهذه العمليات غالبًا. كما تهدف معظم البحوث الصناعية إلى تطوير منتجات جديدة في المدى القريب، في حين تأتي التطورات الأساسية في مجال استخدام الحاسبات من المختبرات الصناعية، مثل زيروكس بارك Zerox PARC ومختبرات بيل Bell Labotation، ومختبرات آي بي إم IBM. وقد قامت شركة مايكروسوفت حديثاً بتكوين فريق بحث رفيع المستوى لهذا الغرض.

إن كثيراً من التقنيات المعروفة التي عملت على جعل مقومات المكتبات الرقمية ذات جدوى كان قد تبناها أناس كانت اهتماماتهم الأولى في مجالات أخرى. فالإنترنت التي لا تقوم بدونها قائمة للمكتبات الرقمية، طُورت أساساً من قبل آربا، والمؤسسة الوطنية للعلوم (NSF). كما أن الويب تطورت في كنف المختبر الأوروبي للفيزياء (سيرن CERN) الذي تلقى دعماً كبيراً من المؤسسة الوطنية للعلوم. كذلك تطور الموزايك MOZAIC، أول متصفحات الويب في كنف مركز استخدام الحاسبات فائقة القوة المُمول من قبل المؤسسة الوطنية للعلوم بجامعة إلينوي، وهناك العديد من مجالات علوم الحاسب، والتي تشمل الشبكات، والنظم الآلية اللامركزية Distributed Computer Systems، والوسائط المتعددة، وتجهيز اللغات الطبيعية، وقواعد البيانات، واسترجاع المعلومات، وتفاعل الإنسان مع الحاسبات، تعد هامة بالنسبة لإدارة المعلومات؛ حيث إن الغالبية العظمى من اتجاهات البحوث في هذه المجالات لها جذورها التي تسبق الاهتمامات الحالية في المكتبات

الرقمية.

وفضلاً عن تمويلها لبحوث معينة، فإن الهيئات الفيدرالية تساعد في تطوير المجتمعات البحثية، كما تقوم في الغالب بتنسيق عمليات نشر التقنيات الحديثة؛ فقد قامت المؤسسة الوطنية للعلوم بدعم لجنة هندسة الإنترنت Internet Engineering Task Force خلال السنوات الأولى لها. كما أن اتحاد نسيج الشبكة العنكبوتية The World Wide Web Consortium، المعتمد أساساً على الدعم الذي يقدمه معهد ماساشوستس للتكنولوجيا، قد دُعم أساساً من قبل الجمعيات الصناعية هذا فضلاً عن تلقيه للدعم المالي من داربا.

ولم تكن المكتبات الرقمية موضوعاً واضحاً على خريطة البحث الفيدرالي حتى التسعينات. وفي عام ١٩٩٢م قامت داربا بتمويل مشروع التقارير الفنية لعلوم الحاسب Computer Science Technical Report Project التي كانت تنسقها مؤسسة مبادرات البحث الوطني Corporation for National Research initiatives، والتي تضم خمس جامعات، هي جامعات: كارنيجي ميلون، وتورنيل، ومعهد إم آي تي MIT، وستانفورد كاليفورنيا في بيركلي، وقد شجع هذا المشروع أقسام علوم الحاسب الآلي ذات المستوى العالي في تلك الجامعات على تطوير برامج البحث في المكتبات الرقمية. وبالرغم من ذلك فلم تظهر تلك المبادرة التي كانت أبرز العلاقات على طريق نشأة المكتبات الرقمية بشكل حقيقي كمجال من مجالات البحث إلا في عام ١٩٩٤م، عندما تبنت كل من المؤسسة الوطنية للعلوم، وداربا، والهيئة الوطنية للطيران والملاحة الجوية مبادرة المكتبات الرقمية.

وقد استطاعت مبادرة المكتبات الرقمية أن تحول أنظار الاهتمام العالمي

نحو البحث في مجال المكتبات الرقمية. وفضلاً عن العمل المحدد الذي قامت بتمويله هذه المبادرة، فقد جسد برنامجها صورة هذا التخصص الناشئ. صحيح أن البحث في مجال المكتبات الرقمية لم يكن جديداً، لكنه كان مجزءاً، حتى أن التسمية ذاتها "المكتبة الرقمية" كانت هشة، وقد استطاعت هذه المبادرة أن تسلط الضوء على المكتبات الرقمية كمجال بحثي تحيط به آفاق التحدي والإنجازات التي تستحق التقدير، كما أنها أثارت الحاجة إلى عقد المؤتمرات، وإلى النشر العلمي، وإنشاء أقسام أكاديمية تجمع كل من لهم اهتمامات مشتركة لعمل الأبحاث في مجال المكتبات الرقمية وإعداد البحوث في هذا المجال. وقد كان هذا التوجه نحو نشأة هذا المجال الجديد هاماً حيث إنه عمل على خلق جو من الثقة اللازمة لمتطلبات البحث العلمي بعيد الأمد.

وبالإضافة إلى ما سبق، استطاعت مبادرة المكتبات الرقمية أن تظهر الفروق بين البحث العلمي، وتنفيذ مشروعات المكتبات الرقمية. فعندما أعلن عن المبادرة اعتقد بعض الناس أن الحكومة الفيدرالية كانت بصدد تقديم مصدر جديد للموارد المالية التي ستتبنى بها المكتبات الرقمية، غير أن المشروعات التي ظهرت كانت حقاً بمثابة مشروعات بحثية. مع أن بعض أنشطة هذه المشروعات تحولت بالفعل إلى تطبيقات عملية، في حين شارك بعضها الآخر في التطورات المرتبطة بالبرمجيات والمكونات المادية للحاسب الآلي، كما كان بعضها الآخر بحوثًا تجريبية حقيقية.

ويثير نشوء المكتبات الرقمية كمجال بحثي أمراً خطيراً، وهو أن الباحثين قد يركزون على القضايا النظرية المثيرة في مجال الحاسبات،

فضلاً عن الجوانب الاقتصادية والاجتماعية والقانونية متناسين أن مجال المكتبات الرقمية مجال تطبيقي في المقام الأول، ويجب أن يحكم على البحث فيه بمدى منفعته، ومن حسن حظ هذا المجال أن هيئات الدعم الفيدرالية كانت لديها الإحاطة بمثل تلك المخاطر، وكانت تحرص على ألا يتعدى التركيز حدود الاحتياجات الفعلية.

لقد قدم الكونجرس الأمريكي دعماً مالياً لكل من داربا وناسا والمؤسسة الوطنية للعلوم (NSF) من أجل تحقيق أهداف محددة، على أن تعمل كل واحدة من هذه الهيئات في إطار محدد ومعروف. ولم يكن لأي من هذه الهيئات مكتبات تعد رسالة أولية لها؛ فأموال مبادرة المكتبات الرقمية جاءت من جراء الدعم المالي الذي رصده الكونجرس لأغراض البحث العلمي في مجال الحاسب الآلي، ونتيجة لذلك فقد ركز الجانب الأول من اتجاهات البحث على الجوانب المرتبطة بالحاسبات الآلية، وكانت هذه الهيئات تدرك أن المكتبات الرقمية أكبر من كونها فرعاً من فروع علوم الحاسب. وفي عام ١٩٩٨م، وعندما بدأت تلك الهيئات في وضع المرحلة الثانية من مبادرة المكتبات الرقمية، بحثت عن شركاء آخرين مهتمين بدعم الأنشطة الموسعة. وقد شملت هذه الأطراف الجديدة المكتبة القومية للطب، وهيئة الوقف الوطني في الإنسانيات، ومكتبة الكونجرس، وقسم التعليم الجامعي بالهيئة الوطنية للعلوم. وتجدر الإشارة إلى أن هذا الكتاب قد تم تأليفه قبل أن يعلن هؤلاء الشركاء ما سيقدمونه من منح في هذا السياق، ولكن يتوقع الجميع رؤية الدعم المالي لمشروعات كبيرة تعكس حالة هذه الهيئات.

وفي سبيل تجسيد كل منها لبرنامجها العملي، قامت هيئات التمويل هذه

بتنظيم سلسلة من ورش عمل، دار بعضها حول موضوعات بحثية محددة، كإدارة الوصول إلى المعلومات، في حين كان للورش الأخرى أهدافها العامة حول تطوير رؤية موحدة للمجال، وتحديد موضوعات بحثية أساسية. وقد كان للمبالغ القليلة من التمويل الفيدرالي المخصص لأغراض تنسيق البحوث حول المكتبات الرقمية أهميتها الحيوية في تجسيدها صورة هذا المجال، وتجدر الإشارة إلى أن كلمة "التنسيق " في هذا الصدد قد حظيت باهتمام أكبر مما حظيت به كلمة "التقييس"، وبالرغم من أن هذه الجهود أحياناً ما تؤدي إلى إنتاج "مواصفات قياسية"، فإن دورها الأساسي انصب على تحفيز إجراء البحوث ورفع معدلات تأثيرها.

ولأن المكتبات الرقمية طورت نفسها كمجال للبحوث الحيوية، فقد انبثقت موضوعات بحثية محددة كبحوث أساسية وينخرط كثير من الباحثين الآن في إعداد هذه البحوث. ويعد الجزء المحوري من هذا الفصل بمثابة نظرة سريعة على الجبهات البحثية الأساسية في مجال المكتبات الرقمية.

اللوحة رقم (٤ - ١)
ائتلاف المعلومات الشبكية
Coalition of Networked Information

يعد ائتلاف المعلومات الشبكية نقطة تحول محورية بالنسبة لآفاق الابتكار بين المكتبيين والناشرين في الولايات المتحدة الأمريكية، وبالرغم من أن الغالبية العظمى ممن يحرصون على حضور اللقاءات نصف السنوية التي ينظمها هذا الائتلاف هم من العاملين في المكتبات الجامعية أو مراكز استخدام الحاسبات الآلية، فإن هذه اللقاءات تجتذب الآخرين كالناشرين

وشركات الحاسبات والمكتبات الوطنية وحكومة الولايات المتحدة، كما شهدت السنوات الأخيرة حضوراً لكثيرين من خارج الولايات المتحدة.

ويعد هذا الائتلاف شريكاً لجمعية المكتبات البحثية والتعليمية Association of Research Libraries and Educause (وهي جمعية لمراكز الحاسبات الآلية الجامعية). وقد تأسس هذا الائتلاف في مارس ١٩٩٠م بهدف تحقيق الأداء العالي للشبكات والحاسبات من أجل تطوير نشاط البحث العلمي وإثراء الإنتاجية الفكرية. ومن بين المئتي مؤسسة الأعضاء في هذا الائتلاف هناك مؤسسات التعليم العالي، وناشرون، ومقدمو خدمات الشبكات، وشركات إنتاج أجهزة الحاسبات والبرمجيات والنظم. وشبكات وتنظيمات مكتبية، بالإضافة إلى مكتبات عامة ومكتبات حكومية.

وفي عام ١٩٩١م، أي قبل عدة سنوات من ظهور متصفحات الويب، اعتبر هذا الائتلاف واحداً من أولى المنظمات التي استحدثت خدمات معلومات رفيعة المستوى على الإنترنت، وفيما يلي قائمة بالأنشطة التي قدمها هذا الائتلاف والتي تم اقتباسها من موقعه على الويب :

- إتاحة خدمات الوصول إلى المعلومات الحكومية الشبكية عبر الإنترنت.
- تقديم خدمات المعلومات لمؤسسات التعليم العالي.
- تقييم كيفية تأثير الشبكات على المؤسسات الأكاديمية.
- توفير ضمانات الثقة والترخيص وإدارة الوصول للمعلومات.

- مراكز التكلفة والمقاييس في سلسلة القيمة للمعلومات الشبكية.
- العمل كاتحاد لعمليات الطباعة الجامعية وتوزيع المعلومات.
- الحصول على رسائل الماجستير والدكتوراة واختزانها.
- عمل لوحات إعلانية إلكترونية للطرق السريعة الرقمية.
- وضع واحد وخمسين سبباً للاستثمار في البنية التحتية للمعلومات الوطنية.
- وضع الاتفاقية المتاحة على الخط المباشر لبيانات شبكة المعلومات الموسعة لمكتب الطباعة الحكومي.
- طرح خدمات الإنسانيات والفنون على الطريق السريع للمعلومات.
- وضع سياسات المعلومات المؤسساتية الموسعة.
- تنظيم ورشة عمل عن ما وراء البيانات Metadata بالاشتراك مع أو سي إل سي OCLC.
- طرح مبادرة وطنية للتراث الثقافي الشبكي.
- تقديم خدمات استكشاف المعلومات الشبكية واسترجاعها.
- استحداث مجتمعات تعليمية جديدة عن طريق الشبكة.
- التقنية والبحث العلمي والإنسانيات.
- منح حقوق الوصول الإلكتروني للمعلومات وبثها.
- تنظيم المؤتمرات الإقليمية.
- تقديم خدمات البحث العلمي من كاليفورنيا على الشبكة.

- تقديم خدمات التعليم والتعلم عن طريق الشبكة.
- حماية الملكية الفكرية في بيئة الوسائط المتعددة الشبكية.
- تقديم خدمات التوصيل للمكتبات العامة.
- التخطيط لملتقى المهنيين في مجال المكتبات وتقنية المعلومات.
- تقديم الموارد الخاصة بالبروتوكول Z39.50 resources.

وتوضح هذه القائمة السابقة الاهتمامات الواسعة لائتلاف المعلومات الشبكية، التي تؤكد التطبيقات العملية للمكتبات الرقمية والمجموعات، والعلاقات بين الناشرين والمكتبات، وقضايا سياسة إتاحة الملكية الفكرية. أما الجانب الذي لم تظهره القائمة فهو الجانب الإنساني لائتلاف المعلومات الشبكية، والذي تبناه كل من مديره التنفيذي السابق باول إيفان بيترز The Evan Peters، وخليفته كليفورد لينش Clifford Lynch، حيث اتصفت اجتماعات ائتلاف المعلومات الشبكية التي جمعت كل الحضور المهتمين من مختلف التخصصات - بالحماس. وتعلم هؤلاء الذين قد تبدو اهتماماتهم متعارضة (متضاربة) كيف يحترم بعضهم بعضًا، وكيف يعملون سوياً. ولقد كان التقدم في السنوات القليلة الماضية يسير بخطى متسارعة بالدرجة التي تجعل من السهولة نسيان ذلك الفراغ الذي ملأه ائتلاف المعلومات الشبكية عن طريق جمع الأشخاص سوياً لمناقشة همومهم المشتركة حول المعلومات الشبكية.

اللوحة رقم (٤ - ٢)

مبادرة المكتبات الرقمية

Digital Libraries Initiative

في عام ١٩٩٤م، قدمت أقسام علوم الحاسب في كل من المؤسسة الوطنية للعلوم، وداربا، وناسا، اعتمادات مالية لستة مشروعات يصل عمر كل منها إلى أربع سنوات، على أن تعمل هذه المشروعات في مجال المكتبات الرقمية. وقد بلغ حجم التمويل الحكومي ٢٤ مليون دولار؛ وقدم هذا التمويل من قبل العديد من الشركاء الخارجيين، وقد توقع لكل واحد من هذه المشروعات الستة أن ينفذ مكتبة رقمية تجريبية، وأن يقوم بأبحاث مرافقة لذلك. وفي الفقرات التالية بيان بهذه المشروعات مع إشارة سريعة لطبيعة كل منها.

- قامت جامعة كاليفورنيا في بيركلي ببناء مجموعة كبيرة من الوثائق عن بيئة كاليفورنيا، وتضم هذه الوثائق خرائط وصوراً، وتقارير حكومية. وتجدر الإشارة هنا إلى البحث الذي تضمن نظام العمل مع الوثائق متعددة القيم multivalent documents (أحد الأساليب الفكرية للتعبير عن الوثائق كطبقات layers للمعلومات)، ونظام تشيشاير الثاني Cheshir II (وهو نظام بحث يمزج بين قوى صيغ المعيار العام للغة العامة الموحدة لتهيئة النصوص (SGML) مع المعلومات في سجلات مارك)، ونظام البحث في مجال تعريف الصور، لكي تظهر تلك السمات المرتبطة على سبيل المثال بالحيوانات والدواجن في الصور.

- ركزت جامعة كاليفورنيا في سانتا باربرا على الخرائط والمعلومات

الجغرافية geospatial، وسميت مجموعاتها هذه بمكتبة الإسكندرية الرقمية Alexandria Digital libraries وقد تضمنت الموضوعات البحثية، وما وراء البيانات عن المعلومات الجغرافية، وواجهات المستفيدين بالنسبة للخرائط المتداخلة، وموجات ضغط الصور وإرسالها wavelets، والأساليب الروائية novel methods لتحليل الكيفية التي يستخدم بها الناس المكتبات.

- قامت جامعة كارنيجي ميلون ببناء مكتبة من مقاطع الفيديو، وتسمى هذه المكتبة بإنفوميديا INFOMEDIA، وقد ركزت البحوث على التجهيز الآلي لاستكشاف المعلومات وعرضها. وكان العمل يتم حول البحث متعدد النماذج multi-model searching - الذي تدمج بموجبه المعلومات التي جمعت من مصادر متعددة ـ، والتعرف الصوتي speech recognition، والتعرف الصوري image recognition، والتصفح المرئي video skimming - الذي يسعى لتقديم ملخص مختصر لمقاطع الفيديو الطويلة ـ .

- قامت جامعة إلينويز بالعمل مع الناشرين على بناء مكتبة فيدرالية federal للدوريات العلمية في العلوم والهندسة. وقد ركزت معظم الجهود على تجهيز الوثائق بالمعيار العام للغة العامة الموحدة لتهيئة النصوص SGML، وقد استخدم هذا المشروع أيضاً الحاسبات فائقة السرعة supercomputing لدراسة مشكلات المعلومات الدلالية في مجموعات ضخمة من الوثائق.

- بني مشروع جامعة ميتشجان على مجموعات المكتبة الرقمية التي أنشأتها مكتبات الجامعة. وفضلاً عن دراسة التطبيقات التربوية، فقد حرص الباحثون على اختبار النماذج الاقتصادية، والأسلوب المعتمد على الوكيل

agent-base approach بالنسبة لعملية التشغيل المتداخل interoperability.

-أما جامعة ستانفورد فقد ركزت على الإنتاج الفكري في علوم الحاسب، وقد سعت إلى إنشاء آلية تعرف بإنفوباص Infobus في مقر المشروع. وهي بمثابة طريقة لدمج الخدمات من مصادر متعددة في مجموعة متماسكة من خدمات المكتبة الرقمية، وقد شمل المشروع موضوعات أخرى منها نمذجة modeling العمليات الاقتصادية في المكتبات الرقمية، وعروض واجهات المستفيدين الروائية novel user interface.

نماذج الكائنات Object Models :

تعد كيفية فهم الكائنات الموجودة في المكتبات الرقمية من أحد الموضوعات البحثية الهامة في هذا المجال؛ فالمكتبات الرقمية تختزن وتبث أية معلومات يمكن أن يتم تمثيلها في شكل رقمي، ونتيجة لذلك، فإن المشكلات البحثية المتصلة بعمليات تمثيل ومعالجة المعلومات متفاوتة ودقيقة.

إن ما يراه المستفيد من المكتبة الرقمية كعمل واحد قد يُمثل في الحاسب الآلي كمجموعة من الملفات وهياكل بيانات موضوعة في صيغ متعددة، وإن العلاقة بين هذه المكونات ورؤية المستفيد للكائن تسمى أحياناً بنموذج الكائن.

وبالنسبة للمستفيد فإن إحدى المقالات المختزنة على خادم الويب ربما تظهر كنص فردي متتابع تصاحبه بعض الرسوم التوضيحية، مع أنه ربما يتم اختزانه كعدة ملفات نصية وعدة صور، وربما بعض البرامج القابلة للتنفيذ Executable Program. كما أن الصورة الفردية قد يتم اختزانها عدة مرات، مرة كصورة أرشيفية عالية الجودة، ومرة أخرى كنسخة متوسطة النقاء

للاستخدام العادي، ومرة ثالثة كنسخة منمنمة صغيرة thumbnails تعطي انطباعاً للصورة بعد حذف الكثير من تفاصيلها. ويمكن الرجوع إلى هذه الصور عن طريق محدد ببليوجرافي واحد، أما بالنسبة للحاسب فتعد المواد بمثابة مجموعة من الملفات الواضحة المعالم. وقد تتوافر للكائن الواحد نفسه عدة إصدارات أو نسخ، وغالباً ما تتوافر للمكتبات الرقمية إصدارات خاصة للمواد التي يتم إعدادها لأغراض البث لعامة المستفيدين. وبعد إطلاق هذه الإصدارات ربما يكون من المطلوب توفير نسخ جديدة لأغراض تصحيح الأخطاء، أو أن تلك المواد يعاد تنظيمها، أو نقلها إلى حاسبات مختلفة، أو قد يتم إضافة صيغ جديدة عند ظهورها.

تعتمد قدرة واجهات المستفيدين وبرامج الحاسبات الأخرى على عرض العمل للمستفيد على البرنامج القادر على فهم كيفية تقريب أو ربط المكونات المختلفة لتشكل كائن مكتبي فردي، على أن تستخدم ما وراء البيانات البنائية Structural Metadata لوصف طبيعة العلاقات بين هذه المكونات. كما تعد لغات التهيئة أحد أساليب تمثيل بنية النص؛ فعلى سبيل المثال، يعد التاج المكتوب هكذا <image> في إحدى الصفحات التي يتم تهيئتها عن طريق لغة تهيئة النصوص الفائقة HTML أحد أشكال ما وراء البيانات البنائية التي تشير إلى موقع إحدى الصور.

ومما تجدر الإشارة إليه إن معظم الجهود المبكرة لما وراء البيانات البنائية قد تمت في سياق مكتبات الصور والموسيقى ولقطات الفيديو المرقمنة، وغيرها من الكائنات الأخرى التي تم تحويلها من وسائط مادية. ويعد مجال الخرائط من موضوعات البحوث الحالية. وبعيداً عن عالم المواد المكتبية التقليدية ـ هناك اهتمام بحثي كذلك بالبيانات الآنية أو الفورية Realtime data

(وهي البيانات التي تم الحصول عليها عن طريق أجهزة الاستشعار عن بعد Remote Sensors)، كذلك هناك اهتمام بالوكلاء المتنقلين mobile agents الذين يتنقلون عبر الشبكات، وبفئات أخرى من الكائنات الرقمية التي ليس لها نظائر مادية. ومن الملاحظ أن كل شكل من هذه الكائنات يثير عدة تساؤلات عن كيفية إنشائه وتخزينه ووصفه والبحث عن المعلومات التي يتضمنها وكيفية تداوله. ومن الواضح أن كل هذه التساؤلات يصعب الإجابة عليها إذا ما نظر إلى كل منها بصورة مستقلة، بل تزداد صعوبة الإجابة عليها إذا ما نظر إليها مجتمعه، ويكمن التحدي في إيجاد نماذج كائنات تدعم المواد المكتبية التي تضم العديد من الأشكال، والتي سوف تمكن المكتبات الرقمية التي نشأت مستقلة بعضها عن بعضٍ من العمل سوياً.

واجهات المستفيدين والتفاعل بين الإنسان والحاسب :

هناك اعتقاد بأن تطوير الكيفية التي يتفاعل بها المستفيدون مع المعلومات المحملة على الحاسبات عملية معقدة، لأنها عملية تنطوي على الجانب الفني أكثر منها موضوعاً للبحث المنهجي، ولكن لحسن الحظ فقد اتضح أنه لا أساس من الصحة لمثل هذه النظرة التشاؤمية؛ حيث استطاعت التطورات التي تشهدها متصفحات الويب أن تسرع من إنجاز البحوث المبتكرة في مجالات كثيرة مثل إمكانية وضع المجموعات المعقدة من المعلومات في صورة مرئية visualization of complex sets of information، ووضع المعلومات التي تشتمل عليها الوثائق في شكل طبقات layering، وعملية التصفح التلقائي automatic skimming لأغراض تلخيص الوثائق، أو إنشاء الروابط create links.

وتعتبر المكتبات الرقمية بالنسبة لمستخدم الحاسب الشخصي جزءاً من بيئة عمل لا أكثر، في حين تنظر بعض البحوث المتصلة بواجهات المستفيدين إلى البيئة الشاملة التي ينتظر منها أن تشتمل على خدمات البريد الإلكتروني وتجهيز النصوص والتطبيقات الخاصة بمجال اهتمام المستخدم. وفضلاً عن ذلك، فمن المحتمل أن تشكل البيئة على كم كبير من المعلومات التي لا تتوافر في شكل رقمي مثل الكتب والمقالات وأشرطة الفيديو والخرائط والصور الفوتوغرافية. وقد يثبت الآن أن مدى قدرة المستفيدين على عمل تعليقات أو وصف الكائنات الرقمية ومعالجتها وإضافتها إلى مجموعاتهم الشخصية تعد مجالاً خصباً للبحث العلمي.

استكشاف المعلومات information discovery :

تقع عمليتا البحث عن المعلومات واسترجاعها في محور اهتمام المكتبات، كما حظيت عملية البحث عن معلومات محددة في مجموعات كبيرة من النصوص والتي تعرف بعملية استرجاع المعلومات باهتمام كبير من جانب علماء الحاسبات. وبالرغم من أهمية عملية التصفح فقد تراجع الاهتمام البحثي بها منذ نشأة الويب، وتجمع المكتبات الرقمية استرجاع المعلومات وتصفحها معاً في قضية عامة هي قضية استكشاف المعلومات، والتي تعني كيفية إيجاد المعلومات، وهناك الآن جهود بحثية كثيرة في هذا المجال، وسوف نشير إلى قليل منها فقط في النقاط التالية :

ما وراء البيانات الوصفية Descriptive Metadata – الفهرسة والتكشيف :

تلجأ الغالبية العظمى من النظم الجيدة لاستكشاف المعلومات إلى

استخدام عمليتـي الفهرسة أو التكشيف لما وراء البيانات التي يقوم بها خبراء فـي هذا الـشأن كالمفهرسـين فـي المكتبات والعاملـين فـي خـدمات التكـشيف والاسـتخلاص. ولكـن لـسوء الحـظ، باتـت عملـية التكـشيف الـيدوية بطيـئة ومكلفة؛ لذا فإن الحجم الضخم من المواد سريعة التغير والمتوقع ظهورها في المكتبات الرقمية سوف تحتاج لوسائل مختلفة [للتعريف بها ووصفها]. وسوف يتم إنتاج بعض من ما وراء البيانات تلقائياً، في حين سيتولى بعض المهنيين المتدربين إنتاج البعض الآخر، هذا في الوقت الذي سيقوم من هم أقل خبرة بإنتاج جزء منها، والبعض الآخر عن طريق إلزام منتجي كل كائن رقمي بتقديم قدر محدود من ما وراء البيانات الوصفية. ومن ثم يمكن تغذية برامج التكشيف الآلية، بما وراء البيانات هذه.

ويـستخدم الـبحث فـي التكـشيف الآلـي بـرامج حاسـوبية للقـيام بالمـسح الـضوئي للكائـنات الـرقمية، واسـتخلاص البـيانات الكـشفية، ومـن ثـم بـناء الكشافات القابلة للبحث. ولعل من أبـرز نـماذج جهـود البحث في هذا المجال بـرامج بحـث الـويب مـثل: ألتافيـستا AltaVista، ولـيكوس Lycos، وإنفوسـيك Infoseek، وإن كانـت الغالبـية العظمـى مـن هـذه البـرامج قـد ظهـرت قبل أن تصبح المكتبات الرقمية مجالاً معروفاً بوقت طويل.

تجهيز اللغة الطبيعية :

لا شـك أن بحـث النص سيكون موفقاً إذا ما تمكن برنامج البحث من فهم بعض بنية لغة هذا النص، وهناك جهود بحثية متصلة بهذا الجانب في مجال اللغويات الحاسوبية، منها ما يتصل بعملية الإعراب أو التحليل اللغوي الآلي automatic parsing، بهدف تحديد التركيبات النحوية، والعمل في مجال علم

الصرف، بهدف معرفة مدى تغير أشكال الكلمة نفسها، وتجمع المعاجم والمكانز. بل إن هناك بعض الجهود البحثية التي تذهب إلى أكثر من ذلك، حيث تحاول التعرف إلى الحالة الموضوعية المتصلة أو المؤثرة في عملية استرجاع المعلومات.

المادة غير النصية :

تلجأ معظم أساليب استكشاف المعلومات إلى استخدام النص، غير أن الباحثين قد حققوا تقدماً بطيئاً فيما يتصل بالبحث عن محتوى محدد في صيغ أخرى مختلفة، ولا تزال عملية التعرف الصوتي في بداية استخدامها في تكشيف برامج الراديو وفي المسارات السمعية لأجهزة الفيديو، كما تعد عملية التعرف الصوري - التي تعني الالتقاط التلقائي للخصائص من واقع الصور - مجالاً نشطاً من مجالات البحث ولكنه مجال غير جاهز حتى الآن ليطرق آفاقاً أوسع.

إدارة المعلومات وحفظها :

لا يزال البحث في مجال إدارة المجموعات في بداية الطريق وإن كان يحظى باهتمام أكثر مما يستحق. وقد سعت المكتبات التقليدية على مر السنوات إلى تطوير أساليب تساعد مجموعات صغيرة نسبياً من الناس في إدارة مجموعات كبيرة من المواد، وفي المقابل كانت المكتبات الرقمية المبكرة تعج بالعمل المكثف جداً في هذا الاتجاه، وفي ظل [حمأة] الاهتمام المتزايد بإنشاء المجموعات الرقمية، أهملت ولفترات طويلة الحاجة إلى تنظيم هذه المواد وحفظها، وقد اتضح الآن مدى أهمية الاهتمام بتلك الحاجات.

التنظيم :

إن تنظيم مجموعات ضخمة من المواد المتاحة على الخط المباشر من المهام المعقدة، ومن الملاحظ أن عديداً من القضايا المشار إليها في هذا الصدد قضايا متشابهة، سواء أكانت المواد الخاضعة للتنظيم دورية إلكترونية، أو موقعًا كبيرًا على الويب، أو مكتبة برمجيات، أو مجموعات خرائط متاحة على الخط المباشر، أو خدمة معلومات كبيرة. وتبدو في الأفق قضيتان هامتان في هذا الصدد هما : كيفية تحميل المعلومات في صيغ مختلفة، وكيفية تنظيم هذه المعلومات لأغراض التخزين والاسترجاع.

وتزداد صعوبة تنظيم المجموعات لأن المعلومات الرقمية عرضة للتغيير. ففي المراحل المبكرة من الطباعة كانت عمليات التصحيح تتم بشكل مستمر، ولذلك كانت كل نسخة من نسخ الكتاب الواحد تبدو كما لو كانت مختلفة بعض الشيء عن النسخ الأخرى. كذلك المعلومات المتاحة على الخط المباشر ممكن أن تتغير بشكل مستمر، وليس من السهولة تتبع التغييرات الطفيفة، أما المجموعات الضخمة فيجب إعادة تنظيمها في جميع الأحوال. ومن الملاحظ أن العديد من الموضوعات البحثية الهامة المرتبطة بعملية التشغيل المتداخل بين المجموعات تكتسب الأهمية نفسها بالنسبة لتنظيم المجموعات، وخاصة تلك البحوث الجارية عن المحددات identifier، وما وراء البيانات Metadata، وأساليب توفير ضمانات الثقة والتحقق من الشخصية authentication المطبقة في عمليات إدارة المجموعات الفردية أو في عمليات التشغيل المتداخل بين المجموعات على حد سواء.

الحفظ الأرشيفي والصيانة Archiving and preservation :

ظهرت عملية الحفظ طويلة الأجل للمواد الرقمية حديثاً كموضوع رئيسي في مجال البحث في إدارة المجموعات، ففي الوقت الذي يمكن فيه تجاهل المواد المحسوسة (كالكتب) لعقود من الزمن، وتظل مع ذلك قابلة للقراءة، فإن الوسائط التي تختزن فيها البيانات الرقمية لا تعمر كثيراً؛ بل إن عمرها غالباً ما يكون قصيراً بشكل مخيف، ولذلك،يجب أن يعاد نسخ البتات bits على وسائط جديدة بشكل دوري. ولتعقيد الأمور، فإن الصيغ التي تختزن فيها المعلومات كثيراً ما يتم استبدالها بإصدارات جديدة. ومن المعروف أن الصيغ التي كانت شائعة الاستخدام في معالجة الكلمات، واختزان الصور قبل عشر سنوات أصبحت الآن غير متوفرة ويصعب استخدامها. ولترجمة المعلومات الأرشيفية، سيحتاج المستفيدون في المستقبل أن يكونوا قادرين على التعرف إلى الصيغ وعلى أساليب عرضها بنجاح.

عمليات التحويل Conversion :

تكشف عملية تحويل المواد المادية التقليدية إلى أشكال رقمية عن درجة الفروق بين الجهود الصغيرة والجهود الكبيرة، وعليه فما هي الطريقة المثلى لتحويل المجموعات الضخمة إلى أشكال رقمية؟ وما العلاقة بين التكلفة والجودة؟ وإلى أي مدى يحتمل أن تكون جهود اليوم مفيدة على المدى البعيد؟ تحتاج المشروعات الصغيرة إلى توفير آلاف قليلة من المواد لاستخدامها في أغراض البحث العلمي، حيث ينظر إلى عملية التحويل وكأنها مصدر إزعاج مؤقت وضروري قبل البدء في البحث الفعلي. وسوف يقوم أعضاء الفريق البحثي بتمرير المواد عبر الماسح الرقمي، ثم تدقيق نتائج المسح تفادياً للأخطاء الواضحة، ومن ثم إنشاء ما وراء البيانات اللازمة لمشروع محدد. لكن

المكتبيين والناشرين عادة ما يقومون بتحويل الملايين من المواد، ومن غير المحتمل أن يكون لأولئك الذين يؤدون هذا العمل دوافع الباحثين العاملين في مشروع صغير ذاتها. وعلاوة على ذلك، وفي إطار جهود تحويل كبيرة، فإن ما وراء البيانات يجب أن تُنشأ بدون معرفة بالاستخدامات طويلة المدى للمعلومات، ولكن ضبط الجودة يجب أن يكون هو القول الفصل لهذه العملية.

وقد سعت بعض المؤسسات إلى تطوير عمليات تحويل فعالة لمجلدات كبيرة من المواد (مع ملاحظة أن الجزء الأكبر من هذا العمل غالباً ما يتم في بلدان تنخفض فيها أجور العمالة)، ومع ذلك فإن لكل مؤسسة من تلك المؤسسات طريقتها الخاصة في التمويل، كما لا يزال هناك تكرار في الأدوات المستخدمة وتدنٍّ لمستوى تبادل الخبرات. ولأغراض تحويل النصوص، وصلت عملية "التعرف إلى الحروف بصرياً OCR - التي تستخدم الحاسب الآلي في التعرف إلى الحروف والكلمات التي تضمها الصفحات - إلى مستوى جيد بما فيه الكفاية. وقد اكتسبت العديد من فرق البحث خبرات هائلة في هذا المجال، غير أن قليلاً من هذه الخبرات هو الذي يحظى بقدر من التنظيم والمشاركة.

مدى قابلية التشغيل المتداخل Interoperability :

من وجهة النظر الحاسوبية، تعد المشكلات الصعبة التي تواجهها المكتبات الرقمية وجوهًا متعددة لتحدٍّ وحيد يتمثل في كيفية تحقيق قابلية الحاسبات للتشغيل المتداخل أو قابليتها للعمل سوياً، والتي تعني كيف يمكن الحصول على تشكيلة واسعة من النظم الآلية التي تعمل بعضها مع بعضٍ. ومن الملاحظ أن عملية التشغيل المتداخل تتراوح في مداها ما بين المستوى التركيبي، حيث

يكون هناك توحد سطحي في الوصول إلى المعلومات وفي عمليات البحث أو الإبحار، مع أنه في حقيقة الأمر اعتماد شبه كامل على إمكانات الذكاء البشري من أجل الترابط، والمستوى الأعمق حيث تتقاسم نظم الحاسب الآلي المنفصلة أدوارها في فهم طبيعة المعلومات.

وهناك الآن في كل أرجاء المعمورة مكتبات رقمية كثيرة يتم إدارة كل منها بشكل مستقل، ولكل منها سياساتها الإدارية المختلفة، ونظمها الآلية المتباينة. كما أن بعض هذه المكتبات حديث النشأة ويعمل بآخر ما وصلت إليه النظم الآلية، في حين أن بعضها الآخر قديم، بل قد بلغ سن التقاعد إن صح التعبير. ومن أحسن ما توصف به تلك النظم القديمة هي أنها نظم موروثة legacy systems [١] مع أن كل واحد من هذه النظم قد أنشئ في الماضي. وما أن يتم التعهد ببناء نظام آلي أو استحداث خدمة جديدة أو منتج جديد، فإن هذا التعهد يصبح أحد العوامل التي توضع في الاعتبار في جميع القرارات المستقبلية. ومن ثم فإن كل نظام آلي يعد نظاماً موروثاً، حتى قبل أن ينتشر كاملاً.

وبالرغم من مدى الارتباط الوثيق بين عملية التشغيل المتداخل والتقييس، فإن الإجراءات الرسمية لاستحداث معايير دولية غالباً ما تكون ـ ولسوء الحظ ـ على عكس ما هو مطلوب لعملية التشغيل المتداخل في المكتبات الرقمية. وليس هذا فقط لأن الإجراءات الرسمية لعملية التقييس عادة ما تكون بطيئة جداً بالنسبة لعالم المكتبات الرقمية المتسارع بشكل كبير، ولكن لأن هذه العملية غالباً ما تشجع على استحداث معايير معقدة بشكل كبير؛ كما أن كثيراً

(١) النظام الموروث: نظام يبقى في حالة استعمال بعد استخدام نظام جديد (المترجمان).

من تلك المعايير لم يخضع للاختبار في الواقع الفعلي، فمن المعروف على مستوى الممارسة أن المعيار لا يكتسب أهميته إلا من خلال تطبيقه واستخدامه على نطاق واسع، وإن كان هناك في بعض الأحيان بعض المعايير تأخذ الصفة الرسمية de facto standard، بسبب تطبيقها واستخدامها من قبل مجموعة مرموقة من الباحثين. ومن أمثلة ذلك معيار بروتوكول التحكم بالإرسال وبروتوكول الإنترنت TCP/IP، كما أن هناك بعض المعايير التي تحظى بالقبول لأن قادة المجتمع قرروا اتباع اتفاقيات معينة certain conventions، وتعد صيغة مارك MARC لسجلات الفهارس مثالاً على ذلك. وعموماً فإن المعايير القياسية المقبولة تستحدث من خلال عمليات تقييس رسمية؛ وخير مثال على ذلك معيار MPEG وهو الصيغة المستخدمة لأغراض ضغط ملفات الفيديو. كما أن بعض المعايير الرسمية هي منتجات محتكرة من قبل شركات كبيرة، ومن أحدث الأمثلة على ذلك برنامج أدوبي لصيغ الوثائق القابلة للنقل Adobe's Portable Document Format. ويلاحظ أن معياري التحكم بالإرسال والإنترنت ومارك قد استحدثا من قبل المجتمعات التي تستخدمها، ثم أصبحا من المعايير الرسمية بعد أن تم تعزيزهما عبر عملية رسمية.

ونشير في الفقرات التالية إلى بعض الجوانب الكثيرة المرتبطة بعملية التشغيل المتداخل :

واجهات المستفيدين :

جرت العادة أن يستخدم المستفيد مجموعات كثيرة من مكتبات رقمية

متعددة، وتسعى جهود التشغيل المتداخل إلى تقديم مواد تلك المجموعات بأسلوب مترابط مع أنه ليس من الضروري إخفاء ما يكمن بين تلك النظم من فوارق. وبالرغم من أن مجموعات الخرائط لا تتطابق في خصائصها مع مجموعة الأعمال الموسيقية، فمن الضروري أن يكون المستفيد قادراً على التنقل بين هاتين المجموعتين بسهولة والبحث فيهما، وألا تعوقه الفوارق الكامنة في نظم الحاسب الآلي أو الخصوصية في كيفية إدارة المجموعات.

التسمية وتحقيق الهوية :

تحتاج المواد في المكتبة الرقمية إلى طريقة ما لتحقيق هويتها، فكل جهاز حاسب متصل بالإنترنت له عنوان واسم مجال. وتقوم المحددات الموحدة لمواقع المصادر (URLs) على الويب بتمديد هذه الأسماء إلى ملفات فردية، ومع ذلك، ليست أسماء المجال هذه ولا المحددات الموحدة لمواقع للمصادر مرضية بما فيه الكفاية. فمواد المكتبات تحتاج إلى محددات تعمل على تحقيق هويتها، وليس إلى الموقع الذي تختزن فيه تلك المواد في وقت معين، وتعرف محددات المواقع المستقلة في بعض الأحيان بالأسماء الموحدة لمواقع المصادر Uniform Resource Name (URNs).

الصيغ :

إن المواد في كل الصيغ الرقمية المعروفة تختزن عادة في المكتبات الرقمية، وقد أوجدت الويب معايير رسمية للتقليل من تعدد هذه الصيغ، ومن أكثر الصيغ تداولاً، صيغة " لغة تهيئة النصوص الفائقة للنصوص البسيطة HTML، وكل من صيغة جي آي إف GIF (١) وجيه بي إي جي JPEG (١)

(١) اختصار لـ Graphic Interchange Format (المترجمان).

لملفات الصور، ومع ذلك يكمن وراء هذه الصيغ الأساسية قدر قليل من الاتفاق عليها. لأن النص يمثل تحدياً خاصاً لعملية التشغيل المتداخل، فمعيار آسكي ASCII الذي ظهر في ثمانينات القرن العشرين الميلادي، بوصفه معيارًا أساسيًا لمجموعة الحروف المستخدمة في الحاسبات، لا يتوفر فيه إلا القليل من الحروف المستخدمة في اللغة الإنجليزية. كما أن نظام اليونيكود UNICODE الذي يظهر الآن معيارًا لمجموعة رموز موسعة تدعم مجموعة كبيرة من النصوص المكتوبة، لم يُدعم حتى الآن من قبل الكثير من النظم الآلية، وعلى الرغم من المناصرة الواسعة التي يلقاها المعيار العام للغة الموحدة العامة لتهيئة النصوص (SGML) المستخدمة في بعض نظم المكتبات الرقمية لغة تهيئة لكل من النصوص وما وراء البيانات، فهي على درجة كبيرة من التعقيد من ناحية والمرونة من ناحية أخرى بشكل يجعل من الصعوبة لها أن تحقق معدلات أداء جيدة لعمليات التشغيل المتداخل. ولعل لغة التهيئة الموسعة XML - التي هي إصدارة مبسطة من اللغة الموحدة العامة لتهيئة النصوص- شائعة الاستخدام- سوف تنجح في وضع جسر لسد الفجوة بين بساطة لغة تهيئة النصوص الفائقة (HTML) والعمومية الكاملة للغة الموحدة العامة لتهيئة النصوص (SGML).

ما وراء البيانات Metadata :

لا يمكن أن ننكر أهمية ما وراء البيانات للكثير من أنشطة المكتبات الرقمية، غير أن أهميتها هذه تزداد بالنسبة لعملية التشغيل المتداخل، وكما سبق أن أشرنا، فإن ما وراء البيانات تنقسم إلى ثلاث فئات: ما وراء البيانات

---

(٢) اختصار لـ Joint photographic Experts Group (المترجمان) .

الوصــفية (وهــي تــستخدم للأغــراض الببليوجــرافية ولأغــراض الــبحث والاسترجاع)، وما وراء البيانات البنائية (التي تربط مختلف الكائنات وأجزاء الكائنات بعضها ببعض)، وما وراء البيانات الإدارية (التي تستخدم لإدارة المجموعات وفي التحكم في الوصول إلى تلك المجموعات). وفيما يتصل بالتشغيل المتداخل، فإن بعض وراء البيانات يجب أن يتم تبادلها بين الحاسبات، وهذا الأمر يتطلب اتفاقاً على مسميات حقول ما وراء البيانات، وعلى الصيغة المستخدمة لترميز هذه الحقول. كما يتطلب الأمر وجود اتفاق معين حول الجوانب الدلالية، ومن الأمثلة البسيطة على أهمية الجوانب الدلالية أن هناك قيمة محدودة لوجود أحد حقول ما وراء البيانات وهو حقل "التاريخ"، وخاصة إذا قامت إحدى المجموعات باستخدامه للدلالة على التاريخ الذي أنشئ فيه أحد الكائنات الرقمية، وعندما تستخدمه مجموعة أخرى للدلالة على التاريخ الذي أضيف فيه الكائن الرقمي إلى المجموعة.

البحث الموزع أو غير المركزي Distributed searching :

كثيراً ما يطمح المستفيدون في إيجاد المعلومات المشتتة عبر العديد من المجموعات المنفصلة بعضها عن بعض، وحيث إن كل مجموعة من هذه المجموعات قد تكون منظمة بطريقة مترابطة، ولكن من المتوقع أن تتفاوت ما وراء البيانات الوصفية لهذه المجموعات، كما ستتفاوت الإمكانات المتاحة للبحث فيها. والأسلوب التقليدي لإيجاد المعلومات من خلال البحث عنها عبر المجموعات المتفاوتة يتمثل في إيجاد اتفاق بين المجموعات على مجموعة معيارية مما وراء البيانات، وأن تدعم هذه المجموعات بروتوكولات البحث ذاتها. وعلى نحو متزايد، يدرك الباحثون في مجال المكتبات الرقمية أن تلك غاية يصعب تحقيقها في الواقع، وأنه من الواجب أن يكون ثمة طريقة تمكن

من بحث المجموعات بشيء من السهولة حتى وإن نظمت مواد تلك المجموعات بأساليب مختلفة.

بروتوكولات الشبكة :

يتطلب نقل المعلومات من حاسب آلي إلى آخر قدراً من قابلية التشغيل المتداخل على مستوى الشبكة، وقد تكفل التبني شبه الدولي لبروتوكولات الإنترنت بحل هذه المشكلة إلى حد كبير، ولكن ما زالت هناك ثغرات في الطريق؛ فعلى سبيل المثال، ليست بروتوكولات الإنترنت هي الوسيلة المثالية في حالة توزيع البيانات التي تبث على نحو مستمر كالمواد السمعية أو المرئية، والتي يجب في حال بثها أن تصل إلى الطرف الآخر بشكل ثابت وعلى مدار فترات زمنية يمكن التنبؤ بها.

بروتوكولات الاسترجاع :

يقوم الحاسب في واحدة من العمليات الأساسية للمكتبات الرقمية، بإرسال رسالة إلى حاسب آخر من أجل استرجاع مواد محددة. وهذه الرسالة يجب أن تنقل عبر بروتوكول معين. وقد يكون هذا البروتوكول مبسطاً إلى حد كبير كما هو الحال مع بروتوكول "نقل النص الفائق HTTP"، أو بروتوكولاً معقداً. ومما تجدر الإشارة إليه أن البروتوكول المثالي يجب أن يدعم كلاً من أمن ضمانات الثقة لكلا الحاسبين، والتساؤلات عالية المستوى الخاصة باستكشاف الموارد التي يقدمها كل حاسب، وأن يدعم كذلك تشكيلة كبيرة من إمكانات البحث والاسترجاع، بالإضافة إلى دعمه لأساليب اختزان النتائج المتوسطة وتعديلها، وواجهات المستفيدين مع الكثير من الصيغ والإجراءات. ويعد البروتوكول Z39.50 من أكثر المحاولات الطموحة التي

سعت إلى تحقيق هذه الأهداف، غير أنه يتعرض لمحنة الآن بسبب تعقده، وعدم وفائه بكل الاحتياجات.

ضمانات الثقة والأمن :

تتمحور العديد من العقبات الكبرى بالنسبة لعملية التشغيل المتداخل على قضايا توفير ضمانات الثقة. وثمة حاجة إلى ثلاث فئات من تلك الضمانات تتمثل الأولى في توفير ضمانات الثقة للمستفيدين، لأن القليل من سبل ضمانات الثقة هو الذي تم تبنيه على نطاق واسع، فالمكتبات الرقمية عادة ما تجبر على منح كل مستفيد اسم مستخدم وكلمة مرور، وتتمثل الفئة الثانية في توفير ضمانات الثقة للحاسبات، فالنظم التي تتداول المعلومات الحساسة، وبوجه خاص التحويلات المالية أو المعلومات السرية، تحتاج إلى معرفة الحاسبات التي تتصل بها أو ترتبط معها عن طريق الشبكة، والأسلوب البسيط في هذا السياق هو الاعتماد على عنوان بروتوكول الإنترنت IP address لكل حاسب، ولكنه أسلوب مفتوح، وعرضة لسوء الاستخدام. أما النوع الثالث من ضمانات الثقة فيتمثل في إيجاد ضمانات الثقة لمواد المكتبة، فالناس بحاجة إلى أن يكونوا على ثقة بأنهم يتلقون إصدارة أصلية من المادة، لا إصدارة معدلة، سواء تم ذلك عن طريق الصدفة أم بصورة متعمدة. وتتوافر الآن بعض سبل ضمانات الثقة الجيدة، غير أنها لم تنتشر على نطاق واسع بالقدر الذي يسمح بدعم إمكانات التشغيل المتداخل الكاملة.

قابلية التشغيل المتداخل التحليلي أو الدلالي Semantic Operability :

يستخدم مصطلح قابلية التشغيل المتداخل الدلالي بمعناه الواسع للإشارة إلى حاجة الحاسبات إلى المشاركة في تفسير دلالي معين للمعلومات التي تتضمنها الرسائل التي تمررها الحاسبات، ويعالج التشغيل المتداخل قضية قدرة المستفيد على الوصول للفئات المتشابهة من الكائنات الرقمية الموزعة عبر مجموعات غير متجانسة، مع محاولة تعويض درجة التفاوت من موقع إلى آخر، وتتبع جهود التشغيل المتداخل الدلالي مجموعة من القضايا البحثية المتعمقة والتي يصل بعضها إلى حد الصعوبة التعجيزية.

وتوفر الويب مستوى أساسياً من التشغيل المتداخل، غير أن بساطة التقنية الأساسية المرتبطة بذلك والتي ساعدت على تقبلها الواسع تحمل في طياتها أيضاً بعض نقاط الضعف؛ فالمحددات الموحدة لمواقع المصادر URLs تعجز عن الاحتفاظ بأسمائها على المدى البعيد، كما أن لغة تهيئة النصوص الفائقة (HTML) مقيدة بدرجة تنوع المعلومات التي يمكن أن تمثلها، كما أن بروتوكول مايم MIME PROTOCOL [(١)]، الذي يحدد نوع كل مادة - جيد حيثما يستخدم، لكن المعلومات المكتبية هي بطبيعتها أكثر بكثير من رؤية بروتوكول مايم لأنواع البيانات، وكذلك فإن واجهات المستفيدين مقيدة ببساطة بروتوكول نقل النصوص الفائقة (HTTP). كذلك أصبح تطوير التوسعات extension لتقنية الويب من الأعمال التجارية الكبيرة، وفي الوقت الذي تؤدي فيه الاحتياجات الفعلية إلى اشتقاق بعض التوسعات، يأتي بعضها الآخر من قبيل التنافس بين الشركات، وتعد لغة جافا JAVA للبرمجة أحد النجاحات اللافتة للنظر في هذا الصدد؛ حيث قدمت إسهاماً كبيراً لواجهات المستفيدين، متغلبة بذلك على العديد من

(١) انظر تعريفًا كاملاً بهذا البروتوكول: ص ٤٦١ وما بعدها (المترجمان).

العقبات التي ترتبط ببروتوكول نقل الملفات الفائقة (HTTP).

وربما يكون من التناقض القول بأن نجاح الويب يعد حاجزاً أمام الجيل القادم من المكتبات الرقمية؛ حيث أصبحت الويب نظاماً موروثاً، وإذا ما أراد الباحثون لأعمالهم أن تلقى قبولاً فيجب أن يقدموا سبيلاً للانتقال من الويب التي نعيشها اليوم إلى خليفتها الجديدة في المستقبل، ويمكن أن تتضح الحاجة الفعلية لمثل هذا السبيل، من الحقيقة التي تقول إن متصفحات الويب الرائدة لا تستطيع دعم الأسماء الموحدة لمواقع المصادر URNs؛ وقد بات ذلك عائقاً في سبيل استخدام الأسماء الموحدة في تحديد هوية مواد المكتبات الرقمية.

ومما تجدر الإشارة إليه أن عملية التشغيل المتداخل من العمليات التي يسهل وصفها ولكن يصعب تحقيقها، وعليه فإن على الباحثين أن يعملوا على تطوير مفاهيم جديدة تكون قادرة على توفير تطورات هائلة ومن السهل تقديمها. كما يجب أن تكون الأساليب الجديدة ذات فعالية عالية، حتى تتغلب على القصور الذي يعم الأساس الذي تم تركيبه أو تهيئته، وأن تكون هذه الأساليب كذلك غير مرتفعة الثمن حتى يمكن تبنيها، وإن التصميم الدقيق لقابلية التوسع extensibility في نظم المكتبات الرقمية سوف يسمح بالتقدم المستمر مع أقل قدر من التعطيل للأساس الذي تم تركيبه.

الحجم Scale :

تصبح عمليتا التشغيل المتداخل وإدارة المجموعات أكثر صعوبة كلما ازداد حجم المكتبة، فالمستفيد قد يجد صعوبة في استخدام الفهرس الخاص بالمنفردات في إحدى المكتبات الكبرى مثل مكتبة جامعة هارفرد أكثر من الصعوبة التي يجدها عند استخدامه لفهرس إحدى مكتبات الكليات الصغيرة،

حيث يكون هناك مداخل قليلة تحت كل عنوان رئيسي. وقد يتفق الكثيرون على أن الويب كلما تطورت فإن برامج التكشيف مثل إنفوسيك Info seek، قد تصبح أقل فائدة، لأنها كثيراً ما تستجيب للاستفسارات المبسطة بالإجابة عنها بمئات من المواد المسترجعة المتشابهة أو ما يعرف بالإصابات hits، وقلما تجد نتائج قيمة بين المكررات والمواد غير الصالحة. ومن الصعوبة إجراء بحث على حجم المكتبات الرقمية دون بناء مكتبات ضخمة. ومن الملاحظ أن البحوث الحالية في هذا الجانب تركز على الجوانب الفنية، وخاصة فيما يتصل بمدى الاعتماد بها على أداء تلك المكتبات أو الوثوق بها.

درجة الثقة أو الموثوقية ومدى القوة Reliability and Robustness :

تعم الأسئلة حول درجة الثقة ومدى قوة الخدمات في مجال المكتبات الرقمية، كما أن درجة تعقد النظم الآلية الكبرى يتجاوز حدود قدراتنا على الفهم الكامل لكيفية تفاعل جميع أجزاء هذه النظم. وفي سياق بعض النظم الضخمة، من المحتم أن تتعطل بعض المكونات أو تغدو خارج الخدمة في أي لحظة، وأحد الأساليب العامة للخروج من هذا المأزق هو توفير نسخ أخرى من البيانات، وغالباً ما يستخدم أسلوب التكرار أو التطابق mirroring، أي تحميل نسخ مكررة من البيانات في عدة مواقع، ولكن لسوء الحظ نادراً ما تكون المواقع المكررة mirror site مطابقة تماماً للأصل بسبب عمليات التأخير التي تفرضها الظروف التي ترافق عملية إعادة نسخ البيانات من مكان لآخر. وعليه يبرز التساؤل التالي: ما الآثار المرتبة على الاسترجاع الموزع أو

اللامركزي إذا لم يكن من الممكن بحث بعض أجزاء المجموعة اليوم، أو إذا استخدمت نسخة احتياطية تحتوي على معلومات قديمة إلى حد ما ؟

إن البحث في مجال معدلات الأداء يعد فرعاً من البحث في مجال شبكات الحاسب الآلي، وليس فرعاً يختص بالمكتبات الرقمية وحدها، ومع أن الإنترنت تكاد تعم جميع أرجاء المعمورة، إلا أن أداءها ليس بالكفاءة نفسها في جميع الدول وفي جميع الأوقات. وتساعد عملية الاختزان المؤقت Caching- أي اختزان نسخ مؤقتة من المعلومات المستخدمة حديثاً، سواء على حاسب المستفيد أو على حاسب خادم آخر- في تحقيق أداءٍ عالٍ عبر الإنترنت واسعة الانتشار على المستوى العالمي، ولكنها تحمل في طياتها الكثير من التعقيدات؛ فماذا يحدث - على سبيل المثال - إذا كانت تلك النسخ المؤقتة متقادمة؟ لاشك أن كل جانب من جوانب الأمن والتحكم في الوصول إلى المعلومات سيكون معقداً، خاصة إذا عرفنا أن اختزان المعلومات يحتمل أن يتم في نظم اختزان مؤقتة غير آمنة حول العالم.

لقد اعتمدت بعض البحوث الهامة عن قياس الأداء على مفهوم "المحلية locality "، ذلك المفهوم الذي يتم بمقتضاه إعادة نسخ المعلومات التي يتم اختيارها، واختزانها في أحد الأماكن المختارة بحكم ما يتوافر لها من إمكانات الاتصالات الجديدة بالإنترنت، فعلى سبيل المثال، تستخدم مكتبة المراجع الفنية الشبكية لعلوم الحاسب the networked computer science technical reference library سلسلة من المناطق zones، بحيث تقوم كل منطقة منها باختزان كل ما هو مطلوب للبحث عن المعلومات ولتحقيق هوية هذه المعلومات، ولا يتم إرسال الرسائل إلى خارج حدود أي من هذه

المناطق إلا بهدف استرجاع الكائنات الرقمية الفعلية من مستودعاتها الأصلية.

قضايا اقتصادية واجتماعية وقانونية :

تتواجد المكتبات الرقمية داخل إطار اجتماعي واقتصادي وقانوني معقد، وللقضايا القانونية أبعادها المحلية والدولية، حيث إنها تمتد عبر عدة فروع من القانون، تشمل حقوق التأليف، والاتصالات، والخصوصية، والأعمال غير الأخلاقية، والقذف، والأمن القومي، والضرائب. ويشمل الجانب الاجتماعي قضايا التأليف، والملكية، واتفاقيات النشر، وضمانات الثقة والتكامل، وهي جميعها ليست مجالاً سهلاً للبحث فيه.

ويعد الجانب الاقتصادي من أصعب التساؤلات التي حظيت بالدراسة، فإذا كانت المكتبات الرقمية حريصة على إدارة مجموعات من المعلومات، فإنها تحتاج إلى مهنيين مؤهلين لإدارة تلك المجموعات، والسؤال هنا، من يدفع أجور هؤلاء؟ والحكمة تقول بأن المستفيدين من تلك المجموعات أو مؤسساتهم سوف يقومون بدفع اشتراكات أو مبالغ مالية نظير كل استخدام لتلك المجموعات. وبناء على ذلك، يمكن القول بأن البحث يتم بناء على أساليب الدفع، وعلى توفير ضمانات الثقة، وعلى أساليب التحكم في استخدام المجموعات. وحالياً أظهرت الجودة العالية للعديد من مواقع الويب المتاحة مجاناً open-access بأن هناك نماذج مالية أخرى. وقد طور الباحثون بعض النظريات الاقتصادية الجيدة، غير أن التطورات الحقيقية في فهم القوى الاقتصادية تأتي ممن ينشؤون المعلومات الرقمية أو يديرونها أو يستخدمونها، ولا شك أن نماذج التسعير ستنقطع عنها صفة الأكاديمية عندما تتسبب

الأخطاء في فقدان الأفراد وظائفهم، أو تخرج المؤسسات عن دائرة أنشطتها.

وأحياناً ترغب المكتبات والناشرون في ضبط عملية إتاحة موادهما والوصول إليها، وربما يحدث ذلك بهدف ضمان عملية الدفع، أو لاستكمال المتطلبات التي يضعها أصحاب حقوق التأليف أو الضوابط التي يضعها المتبرعون، أو استجابة لبعض الاهتمامات الخاصة بموضوع الخصوصية، أو القذف والأعمال المنافية للآداب. ومثل هذه الأساليب تسمى بإدارة الحقوق right management، مع أن المصطلح الأفضل هو إدارة الوصول access management لأن قضايا الوصول أكثر اتساعاً من عملية ضبط حقوق التأليف المبسطة، أو تخليق الإيرادات the generation of revenues، كما أن بعض أساليب إدارة الوصول تتضمن عملية التشفير Encryption، وهذه العملية - أي التشفير - تعد مجالاً معقداً بشكل كبير حيث أصبحت التقنية والقانون والسياسة العامة من الأمور المعقدة بصورة ميؤوس منها.

وقد ثبت أن من الصعوبة بمكان تقييم التأثير الذي تحدثه المكتبات الرقمية، فهل يمكن قياس قيمة المكتبات الرقمية وقيمة البحوث في هذا المجال؟ ولسوء الحظ، وبرغم بعض الجهود النبيلة، ليس من الواضح كيف يمكن الحصول على المعلومات المفيدة؛ فالنتائج المنهجية قليلة ونادرة، والمشكلة معروفة بشكل جيد في بحوث السوق، وهناك أساليب فعالة جداً في التنبؤ بالكيفية التي سوف تؤثر بها التغيرات المتزايدة في المنتجات الحالية، من هذه الأساليب مجموعات التركيز والمسوحات، لكن هذه الأساليب تعتبر أقل فعالية في توقع تأثيرات التغيرات الأساسية، وهذا لا يعني أنه لا حاجة

للمقاييس. ومن غير الممكن تطوير أي نظام كبير، بدون توفير معلومات إدارية كافية مثل: كم عدد المستخدمين لكل خدمة؟ وما تكلفة كل وحدة من وحدات هذه الخدمة؟ وما تكلفة إضافة المواد إلى المجموعات؟ وما حدود التأخير التي يمكن أن تحدث؟ ومن الجدير بالإشارة التأكيد على ضرورة وجود مثل تلك النظم التي تقدم مثل هذه البيانات، وإن كان ذلك يعد من قبيل الممارسة الحاسوبية الجيدة وليس بحثاً بالمعنى الدقيق.

وإذا كنا قد أرجأنا الحديث عن القضايا الاقتصادية والاجتماعية والقانونية إلى نهاية هذا العرض، فليس ذلك لعدم أهميتها وإنما لصعوبة تناولها. وثمة معياران هامان يجب مراعاتهما عند اختيار موضوعات البحث، أولهما إن موضوع البحث يجب أن يكون جديراً بالاهتمام أو يستحق ما يبذل فيه من عناء، وثانيهما أن موضوع البحث يجب أن يكون مجدياً أو ذا فائدة. ويلاحظ أن قيمة المكتبات بالنسبة للعملية التعليمية والبحثية وللصالح العام من غير الممكن إخضاعها للقياس الكمي، وقد أعاقت محاولات قياس الأثر المبكر للمكتبات الرقمية المجموعات غير المكتملة والتقنيات سريعة التغير، والمستفيدون المترقبون لاقتناص الفرص الجديدة، كما أن مقاييس النظم الآلية الجديدة المثيرة تعد سجلاً للتاريخ لا بد منه. كما أنها مهتمة بإدراك الأمور بعد وقوعها ولكنها ذات قيمة محدودة في التخطيط المستقبلي، وإن تحديد قيمة المكتبات الرقمية قد يستمر ليكون موضوعاً للأحكام الرسمية أكثر منه موضوعاً للبحث العلمي.

اتجاهات البحوث حول العالم :

المكتبات الرقمية ظاهرة عالمية، كما أن الإنترنت تسمح للباحثين من كل أرجاء العالم بالتعاون اليومي، فالباحثون من أستراليا ونيوزيلاندا أفادوا كثيراً

بعضهم من بعض، وكلهم مساهم لـه أهميته. كما أن نسيج الشبكة العنكبوتية والويب قد تطورا في سويسرا، وقد قدم مشروع المكتبة الإلكترونية البريطانية دوافع إضافية إلى مجموعة مبادرات المكتبات للعمل مع قضايا النشر الإلكتروني وتوزيع المواد في الأشكال الرقمية. وحديثاً، رعى كل من الاتحاد الأوروبي والمؤسسة الوطنية للعلوم سلسلة من اجتماعات التخطيط المشترك، كما تجلى الجهد الدولي اللافت للنظر في سلسلة ورش العمل عن ما وراء بيانات دبلن كور Dublin Core Metadata، التي سيتم الحديث عنها في الفصل العاشر، وتوضح الروايات المنشورة شهرياً في صحيفة المكتبات الرقمية D. L. magazine عن بحوث المكتبات الرقمية، أنها بحق مجال عالمي؛ فخلال السنوات الثلاثة الأولى كانت المقالات ترد من مؤلفين من أكثر من عشر دول، وكما أن المشروعات الأمريكية الضخمة الممولة تمويلاً كبيراً تعد مهمة، غير أنها ليست كل القصة الكاملة.

الفصل الخامس

الناس والمؤسسات وآفاق التغير

إن قصة المكتبات الرقمية هي إحدى قصص التغير، فالمؤلفون والقراء والمكتبيون والناشرون ومقدمو خدمات المعلومات يتبنون التقنية الحديثة بسرعة واضحة، وهذا ما يغير من طبيعة العلاقات بين الناس.

وإذا كان في كل مؤسسة من يرغب من منسوبيها في استخدام النظم الأكثر تقدماً (حتى لو كانت هذه النظم غير ملائمة)، فإن هناك في المقابل من يأنسون بأداء الأعمال بالأساليب التقليدية حتى لو كانت الأساليب الحديثة هي الأفضل، وهذا ما يعرف في بعض الأحيان "بفجوة الأجيال أو الفجوة بين الأجيال generation gap"، وهي تسمية غير موفقة على أية حال؛ لأن البشر بمختلف أعمارهم يمكن أن يستجيبوا لكل الأفكار الجديدة.

وإذا نحينا ما يفضله الأفراد جانباً، فإننا نرى أن سرعة التغير تتفاوت تفاوتاً كبيراً من مؤسسة لأخرى، ومن تخصص علمي لآخر، فبعض مكتبات الهيئات، المكتبات التابعة لبعض شركات الأدوية، تنفق بالفعل أكثر من نصف ميزانية التزويد فيها على الحصول على المواد والخدمات الإلكترونية، هذا في الوقت الذي ستظل فيه المواد المطبوعة والمخطوطات وغيرها من المواد الأخرى في صلب اهتمام المكتبات المتخصصة في مجال الإنسانيات للمستقبل المنظور.

وربما يعد الاتصال المباشر بواسطة الحاسبات من أكثر التغيرات

الأساسية في هذا الصدد؛ حيث يؤدي هذا النمط من الاتصال إلى تغير سلوكيات كل من منتجي المعلومات والمستفيدين منها؛ فما يتوافر في الحاسبات الشخصية من أدوات يسمح لمن يتمتعون بمهارات محدودة بأداء عمليات كانت تتطلب في يوم من الأيام حرفيين مهرة، كما استطاعت برمجيات معالجة النصوص والنشر المكتبي أن تجعل عملية المراسلات المكتبية تتم بدرجة عالية من الكفاءة، وكأن ذلك أشبه ما يكون بكتاب أحسن إخراجه وتصميمه. كما يمكن الآن إنتاج الرسوم البيانية والتوضيحية وغيرها من المواد الإيضاحية الملونة، مع ملاحظة أنه ليس لدى كل شخص بالطبع الرغبة في تعلم هذه الأساليب. كما أن بعض المواد الرديئة لازالت تنتج، ومع ذلك هناك كثير من الناس ينتجون مواد جميلة وفعالة بدون أية مساعدة من محترفين. ولاشك أن قدراتهم على إنتاج المزيد تؤثر على جميع المهنيين في مجال المعلومات، بما في ذلك الناشرون والمكتبيون والأرشيفيون والمكشفون والمفهرسون ومديرو مواقع الإنترنت.

إن قليلاً من الناس يزعمون أن التقنية الحديثة تقضي على الحاجة إلى إدارة مهنية للمعلومات، وهو زعم ساذج؛ فالناشرون والمكتبيون يقومون بأداء أعمال ووظائف أكثر بكثير من مجرد إدارة مصادر المعلومات التقليدية، فخدمات مثل خدمات التحكيم والتحرير والاستخلاص والتكشيف لا ترتبط بأي تقنية محددة. وعلى الرغم من أن الويب تساعد المستفيدين على معالجة معلوماتهم الخاصة وتحميلها، فإن كثيراً من الناس يسعدون بتلقيهم الدعم من مديري الويب المحترفين؛ ومن ثم فإن الحاجة الكاملة إلى مهنيي المعلومات سوف تستمر وربما تتزايد؛ لأن ممارستهم المحددة تتغير بتغير

التقنية، ومع ذلك فإن الأجيال الجديدة من المؤسسات، وكذا الأجيال الجديدة من المهنيين ـ والتي ستتاح لهم فرص الظهور إلى الوجود، أمر لا يمكن التنبؤ به.

المكتبات الرقمية التي يُنشئها المستفيدون :

لقد أنشئت بعض المكتبات الرقمية التي تعد من أكثر المكتبات الرقمية نجاحاً على أيدي باحثين أو مجموعة من المهنيين، وذلك لحسابهم أو لحساب زملائهم، وبدعم محدود من بعض الناشرين والمكتبيين. وقد تم وصف اثنتين من هذه المكتبات في الفصل الثاني من هذا الكتاب، وهما: أرشيفات المطبوعات الإلكترونية للعلوم الطبيعية في مختبر لوس آلاموس الوطني Physics E-print Archives at the Los Alamos National Laboratory، وسلاسل طلب التعليق عبر الإنترنت Internet RFC series. وتسعى اللوحات الواردة في هذا الجزء من الفصل الحالي إلى وصف ثلاث مكتبات أُخريات هي: مكتبة نت ليب الرياضية The Netlib library of mathematical software ، وأرشيفات البيانات للاتحاد الدولي للبحوث السياسية the data archive of the International Consortium for Political Science Research، ومجموعة بيرسيوس للنصوص الكلاسيكية Perseus collection of classical texts. وتجدر الإشارة إلى أن هذه المكتبات الثلاثة تعد من المكتبات الرقمية الراسخة، كما أنها من المكتبات التي يستخدمها الباحثون بكثافة. وعلى الرغم من وجود مهنيين على رأس العمل بها، فإن قيادتها و معظم موظفيها ينتمون إلى تخصصات متنوعة كالفيزياء والحاسبات الآلية، والرياضيات التطبيقية، والعلوم الاجتماعية والدراسات الكلاسيكية.

إن المكتبات الرقمية التي تُنشئها جماعات من المستفيدين لهي مكتبات لافتة للنظر، لأن خدماتها قد جاءت تلبية لحاجات التخصصات العلمية دون تصورات مسبقة لكيفية إدارة المجموعات في شكلها التقليدي. وعندما يطوّر المنشئون والمستفيدون نظماً محددة، فإنهم يطوّرونها من أجل أعمالهم الخاصة، وإنهم يواجهون أسئلة معتادة عن تنظيم المعلومات واسترجاعها، وضبط الجودة، والمواصفات المعيارية القياسية، والخدمات التي تعد جميعها عصب حياة عالمي النشر والمكتبات. وأحياناً يجدون إجابات جديدة وخلاقة لتساؤلاتهم القديمة هذه.

وإذا كانت المكتبات والمتاحف تؤدي دوراً ذا طبيعة خاصة للمستفيدين في مجالات الإنسانيات،حيث تعمل على توفير المادة الخام التي تعتمد عليها الدراسات في تلك المجالات، فإن المكتبات الرقمية يمكن أن توفر سبلاً أكثر اتساعاً للوصول إلى تلك المواد وبشكل أكثر مما تقدمه المجموعات المادية على الإطلاق. ويمكن القول إن جامعة تتوافر لها مكتبة جيدة قد تحظى بالتفرد والتميّز بخدماتها في مجال الإنسانيات على المستويين التدريسي والبحثي، لكنها لن تكون بحاجة إلى استخدام شامل للعناوين الفريدة. وقد بدأت المكتبة البريطانية في رقمنة تراثها وإتاحته على الإنترنت، ومن أمثلة ذلك، مجموعات مجنا كارتا Magna Carta، ومخطوطات بيوولف Beowulf. وقد كان الوصول إلى تلك المواد في الماضي مقتصراً على الباحثين الذين يزورون المكتبة ويقومون بشراء النسخ أو الطبعات المثيلة facsimile editions باهظة الثمن من تلك الأعمال. أما في المستقبل فسوف يكون بمقدور أي

شخص أن يطلع على نسخ جيدة من تلك الأعمال.

اللوحة رقم (٥ - ١)

مكتبة نت لب الرقمية **Netlib**

نت لب هي مكتبة رقمية ذات برمجية رياضية عالية الدقة، وقد أسس هذه المكتبة في عام ١٩٨٥م كلٌ من جاك دونجارا Jack Dongarra، وإريك جروسي Eric Grosse (اللذان لا يزالان رئيسي تحريرها). والآن يتبنى رعاية شؤون هذه المكتبة اتحاد يضم كل من مختبرات بل .Bell lab بشركة الهاتف والتلغراف الأمريكية AT&T، وجامعة تينسي، ومختبر أواك ريدج الوطني the Oak Ridge National Laboratory، وتتوافر للمكتبة مواقع مطابقة mirror site[١] منتشرة حول العالم. كما أنه يتم تحميل مئات الآلاف من البرمجيات سنوياً، حيث تستخدم هذه البرمجيات في كل فروع البحث العلمي.

ومن الملاحظ أن مكتبة نت لب تخصصت في بادئ الأمر في البرامج المنتجة عن البحث في التحليل الرقمي، وخاصة البرامج المناسبة لأجهزة الحاسبات فائقة السرعة ذات البنيات الفنية الموجهة أو المتوازية vector or parallel architecture. ولكن مجموعاتها تشتمل الآن على أدوات برمجية أخرى، وتقارير وبحوث فنية، وبيانات تقييم الأداء، ومعلومات عن وقائع المؤتمرات واللقاءات العلمية المهنية، وعلى الرغم من أن معظم مواد هذه المكتبة تتاح للعامة مجاناً، فإنها تفرض بعض الرسوم أحياناً مقابل بعض الاستخدامات التجارية.

(١) مواقع تضم النسخ المكررة من المعلومات في أنحاء العالم كافة.

إن التاريخ الفني لمكتبة نت لب يمتد عبر فترة من التطور السريع للإنترنت، فقد بدأت هذه المكتبة كخدمة بريد إلكتروني، ثم قدمت في أوقات مختلفة خدمات لواجهات المستفيدين تعمل في بيئة النوافذ المعروفة بإكس ويندوز (X-Windows) (١) وبروتوكول نقل الملفات مجهولة المصدر FTP، وخدمات الأقراص المدمجة، وخدمات جوفر Gopher، وفي الوقت الحالي تستخدم المكتبة تقنية الويب. ولا يزال فريق الموظفين في هذه المكتبة من بين الرواد في تطوير البنية الفنية المتقدمة الخاصة بتنظيم المواد في المكتبات الرقمية وتخزينها.

ويتم تنظيم المجموعات العلمية في نت لب بطريقة عملية راقية، حيث يفترض أن المستفيدين منها هم من الرياضيين والعلماء ممن لهم معرفة ودراية بالمجال، سيعملون على إدخال البرمجية إلى برامج أجهزتهم الخاصة، بناءً على ذلك يتم ترتيب المجموعات في شكل هرمي، مع برمجيات يتم [تقسيمها] وفقاً للتخصص العلمي أو مجال التطبيقات أو المصدر بحيث يكون لكل مجموعة محررها الخاص بها، ويعمل هؤلاء المحررون على توظيف معلوماتهم ـ كل في مجال تخصصه ـ في اتخاذ القرارات بشأن الطريقة المناسبة للتنظيم. وقد طورت نت لب شكلاً من تسجيلات التكشيف تم إعداده وفقاً لاحتياجاتها الخاصة، كما أن هذه المجموعات تصنف وفقاً لدليل البرامج الرياضية المتاحة Guide to available mathematical software والذي يعد كشافًا متداخلاً Cross-index قام بإعداده "المعهد القومي للمعايير والتقنية".

---

(٢) بيئة في النوافذ أنتجها معهد ماساشوسيتس للتقنية لمحطات التشغيل (المترجمان).

اللوحة رقم (٥ - ٢)

اتحاد الجامعات للبحوث السياسية والاجتماعية

The inter-University Consortium for Political Social Research

اتحاد الجامعات للبحوث السياسية والاجتماعية ICPSR هو منظمة غير ربحية، مقرها جامعة ميتشيجان وتعمل منذ عام ١٩٦٢م، وتسمح المكتبة الرقمية لهذا الاتحاد للباحثين في مجال العلوم الاجتماعية باختزان البيانات التي يقوم بتجميعها وإتاحتها للآخرين للإفادة منها، من منطلق أن البيانات التي يتكلف جمعها مبالغ باهظة تفقد قيمتها ما لم يتم تنظيمها وتوثيقها وإتاحتها للباحثين. وتغطي البيانات التي يتم جمعها مجالات علمية واسعة تشمل علوم: السياسة، والاجتماع، والسكان، والاقتصاد، والتاريخ، والتربية، ودراسة المسنين، والقضاء الجنائي، والصحة العامة، والسياسة الخارجية، والقانون.

ولا يهدف اتحاد الجامعات للبحوث السياسية والاجتماعية إلى الربح، ويتمتع بعضويته المئات من الكليات والجامعات في الولايات المتحدة الأمريكية وجميع دول العالم، وبموجب الاشتراكات السنوية التي تدفعها تلك الجامعات يتاح لأعضاء هيئة التدريس بها ولطلابها إمكانية الوصول إلى مجموعات مكتبته الرقمية. أما الآخرون ممن لا تشترك جامعاتهم في عضوية هذا الاتحاد فيمكنهم الإفادة من مجموعاته مقابل رسوم مالية محددة لذلك.

ويضاف إلى أرشيف هذا الاتحاد ما يقرب من مائتي مجموعة بيانات (أي عدة آلاف من الملفات) سنوياً، وبعض مجموعات البيانات هذه كبيرة جداً. كذلك يتم في هذا الأرشيف اختزان الكثير من الوثائق عن بيانات الأدلة الإرشادية التي تشرح كيفية إعداد تصميم الدراسات والقرارات التي

يتخذها الباحثون الأصليون، والكيفية التي يتم بها جمع البيانات وكيفية تهيئتها، هذا فضلاً عن المعلومات الفنية اللازمة للإفادة من تلك البيانات في بحوث لاحقة. ويتم تنظيم مجموعة هذا الأرشيف بشكل هرمي حسب التخصصات الموضوعية، ولذلك يسهل تصفح هذه المجموعة، كما يتم وصف كل مجموعة بيانات عن طريق عمل سجلات مختصرة تحتوي على بيانات الفهرسة الأساسية بالإضافة إلى مستخلص.

وقد ظل هذا الأرشيف موجوداً عبر عدة أجيال من الحاسبات الآلية، ويمكن الآن استخدام واجهة المستفيد الحالية المصممة لأغراض الويب في هذا الاتحاد لأغراض التصفح والبحث في سجلات الفهرس. كما يتم توزيع مجموعات البيانات عبر الإنترنت من خلال بروتوكول نقل الملفات (FTP)، وإتاحة بعض البيانات المختارة على الأسطوانات المدمجة.

اللوحة رقم (٥-٣)

مكتبة بيرسيوس الرقمية **Perseus**

لقد اضطلع مجموعة من أعضاء هيئة التدريس بالكليات في مجال الإنسانيات ببعض المشروعات الكبرى في مجال المعلومات الإلكترونية، وإن كان هناك من المشروعات الأخرى الكثيرة التي يضطلع بها أشخاص لا ينتمون إلى مهنة محددة ولا يتلقون دعماً من مؤسسة معينة. وقد يحدث في بعض الأحيان أن يتبنى أعضاء هيئة تدريس مبتدئين في بعض الكليات أفكاراً جديدة قد لا تروق لأعضاء هيئة التدريس الكبار فيها، ففي جامعة هارفارد وفي منتصف الثمانينات، ابتكر جريجوري كرين Gregory Crane وهو عضو هيئة تدريس مبتدئ في مجال الدراسات الكلاسيكية مشروع بيرسيوس الذي

يسعى إلى استخدام روابط فائقة للوصول إلى المصادر ذات الصلة بمجال الدراسات الكلاسيكية بدءاً من النصوص والخرائط إلى الأدوات المرجعية مثل المعاجم. وكان يهدف من وراء ذلك بشكل خاص إلى إتاحة الفرصة أمام الطالب العادي لتذوق قصائد الشاعر اليوناني بندار Pindar وبذلك يعد هذا العمل المبكر هو نواة لظهور واحدة من أهم وأكثر المكتبات الرقمية أهمية في مجال الدراسات الكلاسيكية.

تغطي مجموعة مكتبة بيرسيوس حالياً الحقبة الإغريقية الكلاسيكية برمتها، وقد امتدت لتغطي فترات أخرى من التاريخ الإغريقي حتى الحقبة الرومانية وما قبلها. وتشتمل المواد المصدرية للمجموعة على نصوص كلاسيكية (سواء في لغتها الأصلية أو مترجمة)، وعلى صور لكائنات مادية (كفازات الزهور)، هذا بالإضافة إلى مواقع معمارية. ومع ذلك ربما يعد أعظم مصدر هو ذلك الجهد المبذول في بناء المواد، وفي قاعدة البيانات التي تربط فيما بين العناصر.

لقد استطاع التركيز على المحتوى أن يمكن مجموعات مكتبة بيرسيوس من الانتقال عبر عدة أجيال من الحاسبات، كما أن النصوص الكلاسيكية ظلت ثابتة نوعاً ما رغم وجود تغييرات بسيطة في الطبعات الجديدة، أما الأعمال المساندة مثل المعاجم والأطالس فهي أعمال معمرة تعيش طويلاً؛ لذا فإن الجهود التي تركزت على الحصول على نسخ دقيقة من النصوص وإعادة تهيئتها عن طريق المعيار العام للغة (SGML)، وربطها بالأعمال ذات العلاقة بها تعد استثماراً طويل الأجل بحيث تكون أعمارها أطول من أعمار نظم الحاسبات. ومكتبة بيرسيوس التي لم يكن لديها سوى مبرمج

واحد من ضمن موظفيها، فهي تعتمد على أشهر تقنيات الحاسب الملائمة والمتوافرة الآن. وقد كانت من أول من طبق منذ وقت مبكر لنظام أبل هايبركارد Apple Hypercard، كما أنها تنشر أسطوانات مدمجة عالية الجودة، وقد انتقلت بسرعة إلى الويب عندما ظهرت هذه التقنية. ومن البرامج الموسعة الوحيدة التي طورها هذا المشروع تلك الأنظمة المعتمدة على القواعد لأغراض تحليل مورفولوجيا الكلمات اللاتينية والإغريقية المصرفة inflected Greek and Latin words.

إنه لمن الصعوبة التكهن بالأثر بعيد المدى لمكتبة بيرسيوس، غير أن الأهداف طموحة. ففي السنوات الماضية أصبحت الدراسات الأكاديمية في مجال الإنسانيات متزايدة على نحو خفي increasingly esoteric ومنفصلة بعضها عن بعض detached. وهذا أشبه تماماً بالمجال الذي لم يستطع فيه كرين أن يواصل العمل في جامعة هارفارد لأنها لم تعتبر دراسته هذه من الدراسات الجادة، لذا فقد انتقل إلى جامعة توفتس، وقد لا تكون جامعة هارفارد قد فكرت في مكتبة بيرسيوس كدراسة جادة، ولكن ذلك لا يقلل من أهميتها بالطبع. أن نصوص المصادر الإغريقية التي تشتمل على الأربعة ملايين كلمة تشتمل على معظم النصوص الأكثر تداولاً. وعندما لا تكون هناك صور مناسبة لأحواض الزهور، فإنه يعرف عن مكتبة بيرسيوس أخذها لمئات الصور الجديدة؛ وتساعد واجهة المستفيد القارئ على الوصول بسهولة إلى الترجمات والمعاجم مع التركيز على المواد الأصلية. إن مكتبة بيرسيوس التي هي كنز للشخص العادي سيزداد استخدامها من قبل الباحثين الذين سينظرون إليها باعتبارها مصدرًا ممتازًا للدراسات التقليدية. ولعل

أكبر إنجاز حققته مكتبة بيرسيوس يمكن أن يكون تقديمها روائع الإنسانيات للعامة، وتعريف علماء الإنسانيات بأن العمومية والدراسة الأكاديمية الجادة يمكن أن يسيرا جنباً إلى جنب.

دوافع المنشئين والمستفيدين :

المنشئون :

لكي تفهم كيف أن عالم المكتبات والنشر عالم متغير، يتطلب الأمر معرفة الدوافع المختلفة التي تدفع الناس إلى إنتاج المواد وتدفع الآخرين إلى استخدامها أو الإفادة منها.

هناك فهم خاطئ بأن الناس ينتجون المواد أساساً بغرض الكسب أو التملك، ورغم أن كثيراً من الناس يتخذون من عائدات المواد التي يصنعونها مصدراً للعيش، فإن للآخرين أهدافًا أخرى؛ فقد كتب داروين أصل الأجناس بغرض نشر أفكاره، وهكذا كان هدف باين Paine عندما كتب "حقوق الإنسان Rights of Man"، وماركس عندما كتب "رأس المال"، والقديس باول عندما كتب رسالته الإنجيلية إلى الرومان. وقد ألفت في العصور الكلاسيكية بعض الكتب والمخطوطات والصور والأعمال الموسيقية والقصائد تهدف إلى تمجيد بعض السادة وتعظيمهم. و كما تم تشييد كثير من المباني الضخمة في العالم، كأهرامات مصر، والمكتبة الفرنسية، بدافع سيطر على بعض الأفراد وهو الرغبة في تخليد ذكراهم. كما أن كثيراً من الصور الفوتوغرافية والمذكرات اليومية والقصائد والرسائل قد كتبها أصحابها بدافع إسعاد النفس، لكنها تصبح يوماً بعد ذلك من المواد الهامة في المكتبات، وهناك قليل من الأنشطة التي تثمر عن مواد جميلة، كالكتابات الفنية والفنون والموسيقى

والفنون المعمارية. وهناك في المقابل أنشطة قليلة تنتج كثيراً من الأعمال التي تنتفي عنها صفة الجمال، وتحتوي الويب على الكثير من المواد الإعلانية والدعائية، التي سيجد بعضها طريقه إلى المكتبات الرقمية.

وتجدر الإشارة إلى أن العمل الإبداعي قد ينتج عن طريق الصدفة المرتبطة ببعض الأنشطة الأخرى. فالقاضي الذي يطرح الرأي القانوني ينتج بذلك مادة ستكون من مقتنيات المكتبات الرقمية، وكذلك الحال بالنسبة لأمين المتحف الذي يعد فهرساً للمواد المعروضة، أو الباحث الصيدلي الذي يطلب حق الحصول على براءة اختراع. كما أن الهيئات الحكومية كثيراً ما تنتج مواد يستفيد منها عامة الناس، كخرائط الملاحة، وتقارير التنبؤات الجوية، والإحصاءات الرسمية، ونصوص المعاهدات، والاتفاقات التجارية وغيرها. وهناك الآن كثير من المواد المحفوظة في المكتبات ودور المحفوظات أو الأرشيف كان الهدف من إنتاجها تقديم رصد واقعي للأحداث أو القرارات مثل: التقارير القانونية، وسجلات الأبرشيات، والسجلات الحكومية، والمناسبات الرسمية، وصور أحداث الحروب.

إن الأشخاص الذين يقومون بتحويل المواد من شكلها التقليدي إلى شكلها الرقمي يمكن أن نعدهم منشئين أيضاً، مع ملاحظة أن عمليات التحويل هذه تتراوح من الجهد الفردي لنقل قصة مفضلة وإتاحتها عبر الشبكة إلى المشروعات التي يتم بموجبها تحويل ملايين المواد. ويمكن أن يتم العمل الإبداعي الحقيقي بأساليب آلية، كما هو الحال عند التقاط بعض الصور من خلال الأقمار الاصطناعية التي تدور حول الأرض.

وثمة فهم خاطئ آخر مفاده أن للمنشئين والجهات التي يعملون بها الدوافع

نفسها، فهناك بعض الأعمال التي قام بها فريق، وهناك في بعض الأحيان أعمال يقوم بها أفراد، فالفيلم السينمائي لا يمكن أن يقوم به إلا فريق، أما القصيدة الشعرية فلا نعتقد أنها ستكون من عمل الفريق، وعلى ذلك فإن بعض المنشئين كالصحفيين الذين يعملون بشكل حر، والمصورين ومؤلفي الأعمال الموسيقية هم أفراد، ولكن هناك فئة أخرى من المنشئين ينتسبون لمؤسسات، ومن ثم فإن الأعمال التي ينتجونها تعد من صميم أنشطة تلك المؤسسات. وعندما يعمل شخص ما في أحد المؤسسات، فإن صاحب العمل غالباً ما يوجه عملية الإبداع ويمتلك بالتالي نتائجها. وفي الحالات التي من قبيل "العمل مقابل أجر" "Work for hire" عادة ما يكون ثمة خلافات بين دوافع الأفراد العاملين ودوافع الجهات التي يعملون فيها؛ فالمؤسسة التي تنتج فيلماً سينمائياً قد يكون هدفها الربح، في حين يكون هدف المخرج هو اتخاذ هذا العمل فرصة للدفاع عن رأي سياسي معين، هذا في الوقت الذي يكون فيه هدف بطل هذا الفيلم هدفاً فنياً في المقام الأول.

إن المنشئين الذين لا يكون دافعهم الأول هو الربح عادة ما يفيدون من عرض أعمالهم على أوسع نطاق ممكن. مع أن ذلك قد يسبب شيئاً من التوتر بينهم وبين الناشرين الذين يعتمد عملهم في المقام الأول على إتاحة فرص السماح للوصول إلى تلك الأعمال بعد دفع الرسوم لذلك. ولعل الدوريات العلمية أبرز مثال للمواد التي تتناقض فيها دوافع المؤلفين تماماً مع أهداف الناشرين أولئك الذين يسعون إلى تحقيق العائد المادي أساساً، في حين يسعى المؤلفون إلى تسجيل نتائج بحوثهم وإلى تحقيق الشهرة.

المستفيدون :

كما هو الحال بالنسبة للمنشئين، يتفاوت المستفيدون من المكتبات بعضهم عن بعض تفاوتاً كبيراً من حيث اهتماماتهم ومستوياتهم من الخبرة، وتسعى المكتبات العامة في المناطق الحضرية إلى خدمة فئات متفاوتة من المستفيدين. وتعد المكتبة مصدر كتب القراءة الترويحية، ومصدر توظيف بتوفيرها للمعلومات عن فرص العمل، هذا فضلاً عن كونها مصدر معلومات عن مواعيد تسيير الحافلات. وتقدم المكتبة كذلك خدمات الإنترنت التي يتمكن المستفيدون عن طريقها من الوصول إلى مصادر المعلومات الطبية والقانونية، وقد يتوافر فيها شرائط الكاسيت التي تروي قصص الأطفال، ومصادر معلومات مرجعية يستخدمها المؤرخون وغيرهم من الزوار غير المتخصصين أو تستخدم حتى من قبل الخبراء.

وفضلاً عن التفاوت فيما بين الأفراد، فقد تتفاوت احتياجات الفرد الواحد من وقت لآخر، وحتى لو كان هناك تطابق في احتياجات فردين مختلفين فإن استخدامهما للمكتبة يختلف بلا شك؛ حيث يمكن لأحدهما أن يستعين بكثافة بالفهارس والكشافات في سبيل حصوله على المعلومات، في حين يكتفي الآخر في ذلك بالاعتماد المكثف على الروابط والإرجاعات الببليوجرافية Citations. ومن ثم كان لزاماً على مصممي المكتبات الرقمية مراعاة هذا التفاوت؟ وعدم افتراض وحدة أنماط لتلك التوجهات، كما لا يُترك الأمر للنظم التي تفترض أن الكل سواءً في أنماط استخدامه للمكتبة.

ويمكن تمييز فئات عريضة من المستفيدين، وهو أمر على أية حال يسري على المكتبات الرقمية كما يسري على المكتبات التقليدية كذلك، وأول هذه الفئات تضم من يستخدمون المكتبة لأغراض التسلية. وهو يأخذ شكل التصفح

غير المنظم في المكتبة الرقمية، والذي يعرف في الإنجليزية العامية بالقراءة السطحية surfing. أما الفئة الثانية فتضم الذين يرتادون المكتبة بحثاً عن معلومات أولية حول موضوع معين، كالمهندس الذي يبدأ دراسة قضية فنية بقراءة مقالة شاملة survey عن تلك القضية، أو كالسائح الذي يجمع معلومات عن بعض الدول التي يرغب في زيارتها. وفي بعض الأحيان يرغب المستفيد في معرفة حقيقة بسيطة: كأن يريد معرفة نص التعديل الأول للدستور الأمريكي؟ أو درجة ذوبان الرصاص ؟ أو الفريق الرياضي الذي فاز في مباراة الأمس؟ إن بعض هذه الحقائق تقدمها مصادر مرجعية [١] معينة كالخرائط والموسوعات والمعاجم، في حين يتناثر بعضها في ثنايا المجموعات الأخرى. وفي بعض المناسبات يرغب المستفيد في الحصول على معلومات وافية عن موضوع محدد، كالباحث في مجال الطب الذي يبحث عن كل المعلومات المتوافرة عما نشر حول تأثير بعض الأدوية المحددة، أو المحامي الذي يرغب في معرفة كل الحالات المشابهة للحالة التي يدافع عنها.

وفي كثير من تلك المواقف يمكن تلبية حاجات المستفيدين بالاستعانة بمصادر متعددة، فعلى سبيل المثال، يمكن لمجموعة من الخرائط والأطالس أن تقدم المعلومات ذات الطبيعة الجغرافية، هذا في الوقت الذي يتطلب فيه الأمر الوصول لمصادر محددة للحصول على معلومات للدراسة الشاملة لبعض الموضوعات. ويعتبر هذا التمييز ضرورياً إذا ما وضعنا في الاعتبار اقتصاديات المعلومات من ناحية (فالمصادر البديلة للمعلومات

(١) يقسم المكتبيون مصادر المعلومات المرجعية وفقاً لتفاوت الحاجات إلى المعلومات إلى فئتين: مصادر ببليوجرافية، ومصادر غير ببليوجرافية أو حقائقية (المترجمان).

تتنافس في أسعارها)، وعندما يكون المرء بصدد دراسة عملية استرجاع المعلومات حيث يكون للبحث الشامل عن المعلومات أهمية خاصة) من ناحية أخرى.

مهن المعلومات والتغير :

نظراً للتزايد الذي تشهده المعلومات الرقمية وإحلالها محل الأشكال التقليدية، وبحكم اختلاف أعراف المكتبيين عن تقاليد الناشرين بشكل خاص، فإن مهن المعلومات تشهد حالة من التغير، فمن غير المستغرب أن يكون رد فعلهم تجاه التغير مختلفاً.

أمناء المكتبات :

من المفيد لكي نرى كيف يؤثر التغير في المكتبيين أن نقف عند أربعة جوانب كلٍّ على حدة، وهي: مديرو المكتبة، والموظفون الذين في منتصف mid-career librarian حياتهم الوظيفية، وتعليم صغار المكتبيين، والأهمية المتزايدة لوجود اختصاصيين في المكتبات من مجالات أخرى، وخاصة من مجال الحاسبات.

بالنسبة لمديري المكتبات فإنهم يتعرضون للضغوط، فإدارة مكتبة كبرى اعتاد الناس أن ينظروا إليها على أنها عمل جيد ومن يقوم به له مكانة اجتماعية مرموقة، وراتب جيد. وفي الوقت الذي لا يزال فيه الراتب المغري والمكانة الاجتماعية متوافرين، فإن طبيعة العمل نفسه قد تغيرت تغيراً كثيراً، لأن المكتبات الرقمية تحمل تطلعات على المدى البعيد، لكن مشكلاتها كثيرة في المدى القريب. كما أن المكتبات تتعرض لضغوط ارتفاع الأسعار، والمستفيدون التقليديون يطالبون بعدم تقليص الخدمات التقليدية، ومستفيدون

آخرون يطالبون بتوفير سريع للخدمات الرقمية كافة. ويضاف إلى ذلك أن مديري المكتبات غالباً لا يحصلون على الدعم الذي يستحقونه من المسئولين الذين عادة ما يكون تقديرهم لدرجة التغير في واقع المكتبات محدوداً. ونتيجة لذلك نرى في كل عام عدداً كبيراً من مديري المكتبات يقرر عدم الاستمرار أو الانسحاب من هذا العمل الذي تسوده تلك الفوضى، ويسعون إلى إيجاد فرصة عمل أفضل مدعاة للاستقرار.

أما بالنسبة للموظفين الذين في منتصف حياتهم الوظيفية فإنهم يرون في المكتبات الرقمية فرصة وتحدياً في الوقت ذاته؛ فالعجز الفعلي في أمناء المكتبات الكبار senior الذين أصبحوا أكثر راحة بفضل التقنية الجديدة،يمثل فرصة كبيرة للنشطاء والمبدعين من الموظفين الذين يشعرون بتآلف أكثر مع تلك التقنيات. وعلى النقيض من ذلك فإن الموظفين الذين لا يشعرون بذلك التآلف مع تلك التقنيات سوف يتخطاهم الزمن؛ فالتكيف مع التغير التقني أكبر من مجرد إعادة تدريب، صحيح أن التدريب مهم لكنه سيفشل حتماً إذا ركز فقط على استبدال مجموعة مهارات ساكنة Static Skills بمهارات أخرى. فالمكتبات بحاجة إلى أناس يُدركون طبيعة هذا التغير الذي يدور حولهم، وإلى آخرين يحبون البحث، ويكونون منفتحين للأفكار الجديدة.

أما بالنسبة لعملية تعليم صغار المكتبيين فإنه يقع على عاتق كليات المكتبات، حيث إن العمل بالمكتبات يغلب عليه طابع العمل الحرفي. ففي الولايات المتحدة الأمريكية تعد درجة الماجستير من كلية المكتبات مطلبًا للكثير من الأعمال في هذا المجال. ولسنوات عديدة ظلت المناهج في

كليات المكتبات مرتجلة ومركزة على المهارات الأساسية التي يحتاجها أمين المكتبة متوسط الثقافة mid-level librarian، وكانت كلية المكتبات في كثير من الجامعات إحدى أضعف الكليات من الناحية الأكاديمية. ولكن في السنوات الأخيرة أدركت بعض الجامعات أن المكتبات الرقمية يمكن أن تكون فرصًا لشكل جديد من كليات المكتبات بمناهج جديدة وبرامج جادة للبحث العلمي.

وإن المكتبات الرقمية بحاجة إلى بعض الخبراء الذين ليسوا من المكتبيين المحترفين كالمتخصصين في مجال الحاسبات والمحامين. غير أن إقحام مثل هؤلاء الخبراء في عالم المكتبات المنظم high structured world of libraries قد ينطوي على بعض الإشكاليات. فالمكتبيون مترددون في توظيف من هم من خارج مجالهم، وهم عندما يحتاجون مرشحين لشغل وظائف جديدة يضعون قيوداً على المرشحين بأن يكونوا حاصلين على درجات جامعية في علم المكتبات. مع أن قليلاً من الشباب الذين يقومون بإنشاء مكتبات رقمية هم من سبق لهم أن رأوا كليات مكتبات في حياتهم الوظيفية.

ومع مقارنة مهنة المكتبات مع المهن الأخرى، يلاحظ تدني الرواتب التي تدفع للموظفين في منتصف حياتهم العملية، فالعمل الراقي يحتاج إلى أشخاص راقين Top class people [(١)] . أما بالنسبة للمكتبات الرقمية فالعاملون فيها يحظون بتقدير عال في سوق العمل. وهناك إشكالية في أن

(١) عبارة غير موفقة من المؤلف، لعل الأنسب لها: أشخاص مؤهلين أو أكفاء أو مميزين أو قياديين (المترجمان).

يُدفع للمبرمج أكثر مما يُدفع لرئيس القسم، ولكن الإشكالية الأكبر هي أن ترى مكتبة جيدة تتدهور بسبب ضعف مستوى موظفي الحاسبات. ولعل جزءًا من النجاح الذي حققه مشروع ميركوري في جامعة كارنيجي ميلون يُعزا إلى أن الموظفين الفنيين العاملين في المشروع كانوا من الناحية الإدارية يعملون في مركز الحاسب الآلي بالجامعة. وبالرغم من أن مكاتبهم كانت في المكتبة، إلا أن ولاءهم كان للمكتبات الرقمية، ومع ذلك كانوا يعملون تحت إشراف مديرين فنيين، وكانوا يعملون في ساعات غير منتظمة، وكان لديهم أفضل الأجهزة، كما كانوا يتقاضون الرواتب نفسها التي يتقاضاها موظفو الحاسب الآخرين. وختاماً إنها قليلة تلك المكتبات التي تتسم بمثل هذه المرونة.

اللوحة رقم (٥ - ٤)

مدرسة تايسر الصيفية **The Ticer summer school**

في عام ١٩٩٦م قدمت جامعة تلبيرج Tilburg الهولندية دورة صيفية لمدة أسبوعين تهدف إلى تثقيف كبيري المكتبيين بمجال المكتبات الرقمية، وقد حقق هذا البرنامج الذي عرف باسم تايسر Ticer ناجحاً سريعاً بحيث أصبح يتم حجزه في كل عام. وقد كان طلابه في أول الأمر من شمالي أوربا، ولكن منذ عام ١٩٩٨م التحق به طلاب من ماليزيا، واليابان، والهند، وكوريا الجنوبية ودول أخرى، وكان معظمهم من كبار الموظفين العاملين في مكتبات أكاديمية أو متخصصة.

لقد كانت جامعة تلبيرج رائدة في تطبيق فكرة المكتبة الرقمية منذ سنوات، وتعكس هذه الدورة التي تقدمها الجامعة وحدة الرؤية بين التخطيط

الاستراتيجي والتطبيق العملي والذي عُدَّ من أحدى علامات بروز جهود هذه الجامعة. وكان معظم المحاضرين في هذه الدورة الصيفية من موظفي مكتبات الجامعة أو مراكز خدمات الحاسب الآلي فيها، بالإضافة إلى بعض الأساتذة الزائرين الذين يضفون على الدورة مزيداً من العمق والرؤية الصائبة. ويعتبر كتاب المحاضرات السميك الذي يحصل عليه الطلاب الملتحقين بالدورة مصدراً رئيسياً للمعلومات.

وقد أبرز النجاح الذي حققه برنامج تايسر هذا مزايا أخرى، حيث كانت التكاليف منخفضة إلى حد ما، برغم توفير خدمات عالية المستوى، يضاف إلى ذلك وجود برنامج اجتماعي يعمل على تعزيز العلاقة بين الذين يلتحقون بهذا البرنامج من كل أنحاء العالم والبالغ عددهم خمسين شخصاً، إذ يعيشون ويعملون بعضهم مع بعض. ولبرنامج تايسر علاقة قوية بالناشر إلسفير الهولندي المعروف، حيث يحضر عدد كبير من موظفي هذا الناشر من جميع أنحاء العالم تلك الدورة طلابًا، في حين يقدم بعض من كبار موظفيه بعض المحاضرات.

إن تجربة تايسر تعكس إلى أي مدى أصبح مجال المكتبات الرقمية مجالاً عالمياً، هذا فضلاً عن أنه يعكس مكانة اللغة الإنجليزية كذلك. ومع أن هذا البرنامج يقدم أساساً باللغة الإنجليزية غير أن المواد الإعلانية لا تذكر هذه الحقيقة، ومع ذلك فلا يزال عدد الطلاب القادمين من الدول الناطقة بالإنجليزية قليلاً.

اللوحة رقم (٥ - ٥)

كلية إدارة ونظم المعلومات في بيركلي

مع بداية التسعينات من القرن العشرين الميلادي أعادت كثير من الجامعات الكبرى النظر في كليات المكتبات فيها، فبينما كانت الكليات المتخصصة في الحقوق وإدارة الأعمال بتلك الجامعات تجتذب الطلاب وأعضاء هيئة التدريس المرموقين والمساهمين في تطوير البحوث فيها وتحقيق الأرباح لها، كانت كليات المكتبات تتهاوى وتتدهور. وكان أعضاء هيئة التدريس فيها من أقل أعضاء هيئة التدريس إنتاجية وأقلهم رواتب كذلك. والأسوأ من ذلك كله أن البرامج التعليمية فيها لم تكن قادرة على تخريج القادة الذين تحتاجهم المكتبات الرقمية الجديدة وصناعة النشر الإلكتروني.

وقد قامت جامعة كولومبيا بإغلاق كلية المكتبات فيها، حيث اعتبر أن برنامجها الأكاديمي أقل قيمة من المباني التي تشغلها الكلية، كذلك قامت جامعة شيكاجو بإغلاق كلية المكتبات فيها، أما جامعتا كاليفورنيا في بيركلي وميتشيجان فقد اختارتا الطريق الآخر وهو إعادة تطوير هذه الكليات، وربما كان التغير الأكبر هو ما شهدته جامعة كاليفورنيا في بيركلي.

إن كثيراً من الجامعات تدعي أن القرارات الأكاديمية فيها يتخذها أعضاء هيئة التدريس، ولكن ذلك ليس سوى ادعاء في كثير من الجامعات، أما في بيركلي فيتمتع المجتمع الأكاديمي بصلاحيات كبيرة. وقد كان التوجه الأول في بيركلي أن تحذو حذو كولومبيا وتغلق كلية المكتبات، ولكن كانت هناك معارضة قوية جعلت الجامعة تعدل عن رأيها، وتكّون مجموعة تخطيط عالية المستوى. وجاء تقرير هذه المجموعة تحت عنوان "اقتراح لكلية إدارة ونظم المعلومات "، ونشر في ديسمبر من عام ١٩٩٣م، ويمكن

القول وبالبيروقراطية التقليدية، أن التقرير نص على تفكيك كلية المكتبات القائمة وإنشاء كلية جديدة على حطامها.

وعادة ما تنظر الجامعات الجيدة إلى عاملين عند وضع برامج أكاديمية جديدة، يتمثل الأول في طبيعة المضمون الأكاديمي؛ بمعنى هل لهذا المجال مضمون فكري عميق يجتذب أفضل أعضاء هيئة التدريس الذين يسهمون في رفع سمعة الجامعة؟، أما العامل الثاني فيتمثل في التعليم، بمعنى هل سيجذب البرنامج الطلاب القادرين الذين سوف يتخرجون فيما بعد ويصبحون قادة؟ وقد أجاب تقرير مجموعة التخطيط في بيركلي على هذين السؤالين بشرط بالإيجاب استكمال بعض المعايير المحددة.

وفي بداية الأمر رفض التقرير صراحة مقررات كلية المكتبات التقليدية، وكذلك درجة الماجستير في علم المكتبات. وبصراحة شديدة أورد التقرير ما يلي: "إن الدرجة الجامعية التي يمنحها هذا البرنامج... لم تصمم لتتوافق مع متطلبات جمعية المكتبات الأمريكية، وإنما تمثل نموذجاً لتطوير معيار منح شهادات في مجال ناشئ مركز عليه".

وتم بعد ذلك قبول التقرير، وبدأت الكلية في استقطاب أعضاء هيئة التدريس والطلاب. وكان التركيز منذ البداية على الدراسات ذات الطابع متعدد التخصصات interdisciplinary، وكان تقرير مجموعة التخطيط ملیئاً بالإشارات إلى البرامج المشتركة مع كليات وأقسام أخرى. كما أعلن البرنامج عن درجة الماجستير الجديدة مشيراً إلى أنها: "جوانب من الدراسات في مجالات الحاسب الآلي والعلوم المعرفية وعلم النفس والاجتماع والاقتصاد وإدارة الأعمال والقانون ودراسات المكتبات والمعلومات والاتصالات". ولم يكن يشترط على

الطلاب الراغبين في الالتحاق بهذا البرنامج سوى التمتع بالمهارات العالية في الحاسبات الآلية.

وكان (تعيين) هيئة التدريس العاملين في هذه الكلية متنوعاً كذلك، حيث كان عميدها الأول، هال فاريان Hal Varian، رجل اقتصاد، كما كان من ضمن أعضاء هيئة التدريس بها علماء متخصصون في قانون حقوق التأليف، بالإضافة إلى الرئيس السابق لقسم علوم الحاسب بالجامعة. وكان لكثير من أعضاء هيئة التدريس بها مسؤوليات متعددة ومشتركة كالتدريس وعمل الأبحاث العلمية،كل في مجاله التقليدي.

ولا يزال الوقت مبكراً لكي نحكم على نجاح هذه الكلية الجديدة، أو على الجهود المماثلة في جامعة ميتشيجان، غير أن المؤشرات الأولية مبشرة، فقد لبت الكلية المتطلبات الأساسية لتعيين أعضاء هيئة التدريس ولاختيار الطلاب. كما أن برامج البحث فيها تطورت بسرعة، وحصلت على منح خارجية كبيرة للبحث في الموضوعات الهامة. وبعد سنوات قليلة من الآن ستظهر طلائع الخريجين من هذه الكلية، والذين سيكون لهم دور القيادة لجيل قادم من مهنيي المعلومات.

الناشرون :

إن التغيرات التي طرأت على عالم النشر تغيرات جذرية كما هو الحال بالنسبة لتلك التي طرأت على عالم المكتبات؛ فمنذ اختراع الطباعة في أوربا في القرن الخامس عشر وعالم النشر يشهد تغيرات بطيئة، واستمر عمل الناشر تقليدياً متمثلاً في إنتاج المواد المطبوعة وتوزيعها. واليوم ما تزال صناعة النشر تحصل على معظم عائداتها من مبيعات المواد المطبوعة

كالكتب والمجلات والجرائد، ثم تلتها أشرطة الفيديو، فالأقراص المدمجة، ومع ذلك يرى كثير من الناشرين أن النمو المستقبلي لصناعتهم سوف يكون في مجال النشر الإلكتروني.

إن صناعة النشر هي خليط من الصناعات الصغيرة وإدارة الأعمال الكبرى؛ فدور النشر الكبرى مثل تايم وورنر Time Warner، وريد إلسفير، ومجموعة تومسون Tomson group، يصنفون على أنهم من كبريات دور النشر على مستوى العالم وأكثرها ربحاً، هذا في الوقت الذي يتواجد فيه أكثر من ٥٠٠٠٠ ناشر في الولايات المتحدة الأمريكية ينشر كل منهم أقل من عشرة عناوين في السنة. أما النشر الأكاديمي فله خاصية غريبة وهي أن بعض دور النشر التي تصنف في هذه الفئة تعد مؤسسات تجارية، وهدفها الأساسي هو تحقيق الربح، في حين يدخل البعض الآخر في المؤسسات التي لا تهدف إلى الربح، كالجمعيات العلمية ومطابع الجامعات، وينصب اهتمامها في المقام الأول على دعم البحث العلمي.

إن نجاح أي دار نشر يعتمد على المحررين الذين يضطلعون بمهام اختيار ما ينبغي نشره، والعمل مع المبدعين، والإشراف أحياناً على بعض الأعمال في مراحل إنتاجها. وبالرغم من أن عملية النشر تعد عملاً تجارياً، فإن كثيراً ممن يعملون فيها لا يكون هدفهم الحصول على عمل تجاري بقدر شغفهم بمضمون المواد التي ينشرونها.

وبشكل عام يلجأ الناشرون إلى استخدام المقاولين أو الوكلاء contractor لإدارة أعمال الطباعة مع أن هناك دور نشر قليلة مثل وست ببلشنج West

publishing (وهي من أكبر دور النشر القانونية) تدير مطابعها بنفسها. غير أن معظم عمليات الطباعة يتولاها متخصصون في هذا المجال، كما يتولى عمليات التحرير الكاملة للمخطوطات التي ستنشر محررون متعاونون freelance، وتباع الكتب التي تنشرها هذه الدار عبر منافذ بيع الكتب، في حين توزع المجلات من خلال منافذ الاشتراكات. ولعل هذا الاستخدام للوكلاء يمنح الناشرين مرونة لا يمكن أن تتوافر لو أن كل شيء كان يتم تنفيذه داخلياً في دار النشر. وعندما تحول الناشر إلسفير إلى مجال إنتاج المجلات الإلكترونية مستخدماً المعيار العام للغة كتابة الوثائق [u1] (SGML)، كان بمقدوره سحب الاتفاقيات من تلك المطابع التي لم تكن تستعد للتحول إلى استخدام المعيار العام للغة كتابة الوثائق [u2] (SGML).

إن موجة الاندماج التي حدثت مؤخراً في صناعة النشر قد أوجدت بضع شركات كبرى لديها الثروة اللازمة لدعم مشروعات التحسيب الضخمة، وبرغم أهمية الحجم إلا أننا مازلنا في انتظار معرفة عائده في مجال النشر الإلكتروني. وإن تقنية الويب تعني أنه بوسع الشركات الصغيرة الانتقال السريع إلى مجالات جديدة، وربما تتميز هذه الشركات الصغيرة في بعض الأحيان بقدرتها على إقامة علاقات قوية بين المحررين والمؤلفين. وإن كان بعض المراقبين يرى أن بيع دار نشر ويست ببلشنج West Publishing والتي كانت تابعة للقطاع الخاص، كان سببه الخوف من أن النشر الإلكتروني قد يضعف من هيمنة تلك الشركة على سوق النشر القانونية.

إن قليلاً من المطابع الجامعية، ومنها مطبعة جامعتي أكسفورد وشيكاجو، تتشابه في أمور كبيرة في علاقتها مع الناشرين التجاريين في

نشر الأعمال الأكاديمية، فهي تنشر أعداداً كثيرة من الكتب الثقافية العامة والمراجع والكتب العلمية، وهذه المطابع تعمل على أساس مالي قوي، كما أنها تعطي الأولوية للمؤلفين من جامعاتها. ولمطابع جامعية أخرى دور آخر يتمثل في دعم البحوث العلمية عن طريق نشر الأعمال العلمية (مثل الكتب أحادية الموضوع الشائعة في مجال الإنسانيات) والتي قد لا تكون رائجة بحيث لا يباع منها إلا المئات من النسخ،ومثل هذه الكتب لا تلقى اهتماماً من دور النشر التجارية. ومن الملاحظ أن مطابع هذه الجامعات تعمل في ظروف مالية صعبة، ففي السنوات الأخيرة، ومع إقدام الجامعات على تقليص ميزانياتها، فإن دعمها للمطابع بدأ يقل، بل قد انقطع تماماً في بعض الحالات.

اختصاصيو الحاسبات ومديرو مواقع الويب :

لاختصاصيي الحاسبات مكانة حيوية في عالم المكتبات الرقمية كما هو الحال بالنسبة للمكتبيين والناشرين، غير أن الفوارق الثقافية والحواجز التي تعوق التعاون فيما بينهم كبيرة؛ فكثير من أفراد مجتمع الإنترنت (١) ترعرعوا في مجال الحاسبات، ولهم أساليب عمل غير مألوفة للآخرين. وكما هو الحال في كل تخصص علمي، يدرك بعض أبنائه التفاصيل التقنية كافة المرتبطة بهذا التخصص والتي لا يرى غيرها، لكنهم غالباً ما يجدون صعوبة حقيقية في وصف تلك القضايا التقنية لغير المتخصصين. بل إن بعضهم يتعمد الغموض، ويعتبرون هذه المعرفة لغزًا، وهو توجه نادر

(١) بالطبع يقصد المؤلف بمجتمع الإنترنت internet community مطوري الإنترنت وليس مستخدميها (المترجمان).

لحسن الحظ؛ فالغالبية العظمى من الناس يتطلعون إلى أن يطلّع الآخرون على ما يقومون به من أعمال، لكن القضية أن الفنيين عادة ما يعانون من وصف التقنية لغير المتخصصين.

ولمجتمع الإنترنت مؤسسوه من جيل المتخصصين في الحاسبات الآلية الذين ترعرعوا في نظام التشغيل يونكس UNIX، هذا النظام الذي تحقق له النجاح الكبير بفضل طبيعته المرنة. ولقد تقاسم هؤلاء الثلاثة - اليونيكس، والإنترنت، ونسيج العنكبوت العالمي (WWW) - الجهود واكتشفوا أن ذلك يدفع كلاً منهم إلى مزيد من الإنتاجية، ولعل تقدير كل منهم لروح الانفتاحية openness هذه كان أساساً لفهم مدى إدراك علماء الحاسبات لتطور المكتبات الرقمية، وكيف أن توجههم قد يصبح حلماً أشبه بحلم المدينة الفاضلة، وقليل هم المثاليون الذين يدافعون عن مجتمع معلومات حر غير محكوم بأية قيود أو رقابة، ولكن تلك وجهة نظر متطرفة. إن الغالبية العظمى من أفراد مجتمع الإنترنت لديهم حب طبيعي للمال، حيث إن العمل التجاري هو جزء من التقليد، وعليه تنشأ شركات جديدة باستمرار.

وحتى في عالم الحاسبات الذي يشهد تغيرات متسارعة، فإن مهنة مدير موقع الويب Webmaster لم يكن لها ما يوازيها. وفي بداية عام ١٩٩٤م لم تكن الويب قد عرفت تماماً، ولكن في صيف ١٩٩٥م حضر أكثر من ألف شخص أول لقاء لمديري الويب الفيدراليين، وهي مجموعة غير رسمية من الأشخاص الذين يقومون بصيانة مواقع الحكومة الأمريكية على الويب. وبين عشية وضحاها ظهر مصطلح "مديرو مواقع الويب" وأصبح متداولاً بين الناس، لدرجة أن محاولات إيجاد بديل عنه باءت كلها بالفشل (وذلك

على اعتبار أن كلمة master تمثل الرجال فقط)، ولكن يجب أن يفهم أن مصطلح مديري الموقع ينسحب على الجنسين، الرجال والنساء معاً.

إن مدير موقع الويب يمكن أن يكون ناشراً أو مكتبيًا، أو خبيرًا في الحاسب الآلي، أو مصممًا، فهذا العمل يتطلب مجموعة من الخبرات، تشمل إلى جانب مهارات النشر التقليدية، مهارات أخرى في الاختيار والتحرير، ومهارات في تصميم واجهات المستفيدين، ومهارات تشغيل نظم الحاسبات الآلية عالية الأداء. فالموقع على الويب لشركة معينة هو الوجه الذي تُرى به تلك الشركة أمام العالم؛ وعليه فإن جودة الرسومات، والأسلوب الذي تقدم به المواد على الموقع لا تقل أهمية عن مادة العلاقات العامة التي تقوم المنظمة ببثها. وفي مؤسسة CNRI تصف مديرة الموقع نفسها - مازحة - بأنها هي قسم الفن the Art department. وحيث يعمل أمين المكتبة المحترف الذي يملك المهارات الفنية العالية، تقضي هي معظم وقتها في أعمال، يمكن تسميتها في سياقات أخرى بالتصميم الجرافيكي (Graphic Design).

وفي بعض المؤسسات يقوم "مدير موقع الويب" باختيار أو إنشاء المواد التي تظهر على موقع الويب، والأمر الأكثر شيوعاً هو أن يتم إعداد هذه المواد من قبل أفراد أو مجموعات من داخل المنظمة. وتتركز وظيفة مدير الموقع على تحرير هذه المواد الفردية وتشكيلها وتنظميها داخل الموقع. ولذلك فإن المظهر العام للموقع يكون من إعداد مدير الموقع، وليس المواد الفردية. وعلى سبيل المثال تقوم مؤسسة CNRI بإدارة موقع لجنة العمل الخاصة بهندسة الإنترنت Internet Engineering Task force (IETF). في حين تتولى سكرتارية اللجنة التحكم بالمحتوى، ويسهم مدير الموقع بتصميم الصفحة

الرئيسية للموقع ووضع الروابط بين الصفحة الرئيسية والمواد الأخرى على موقع الويب.

يتمتع مدير الموقع بمهارات حاسوبية متنوعة، تعتمد على مدى حجم الموقع وصعوبته. كما أن بعض الأساليب المستخدمة في تصميم واجهة المستفيدين مثل لغة جافا تحتاج إلى مهارات المبرمجين. كما أن بعض مواقع الويب تكون كبيرة ومتسعة ومعلوماتها متكررة على حاسبات آلية ذات مواصفات عالية موزعة حول العالم. ويعتبر موقع شبكة أخبار كيبل Cable News Network من المواقع البسيطة المعروفة حيث يستقبل أكثر من ١٠٠ مليون عملية بحث hits يومياً؛ لذا فإن عملية تصميم ومراقبة هذه الأنظمة الكبيرة وتوصيلات شبكتها تحتاج إلى مهارات العمل في مجال الحاسب الآلي. فإذا كان موقع الويب يتعامل مع التحويلات التجارية فلابد أن يكون لدى مدير الموقع الخبرة في مجال أمن الشبكات كذلك.

يتوفر لدى كثير من المنظمات التي تسهم وتشارك في مجال المكتبات الرقمية قسم خاص للحاسب الآلي. حيث يستطيع مديرو مواقع الويب أن يثقوا بهم في القيام بالمهام الفنية العرضية مثل إعداد خادم الويب أو إعداد الكشاف القابل للبحث على الموقع. وفي بعض المؤسسات الأخرى يتوجب على مدير الموقع القيام بجميع تلك المهام بنفسه، كما يحدث أن يكون هناك كثير من مواقع الويب التي تخدم مؤسسات لا يتوافر لها موظفون متخصصون في الحاسب. ومن ثم تقوم ثلة كبيرة من الشركات بتقديم خدمات الويب لمثل تلك المنظمات، من تصميم صفحات الويب، وتشغيل الحاسبات الخادمة، والقيام بالمهام الإدارية كتسجيل أسماء المجالات.

أشكال جديدة من المؤسسات :

الاتحادات Consortia :

إن إدارة مجموعات كبيرة متاحة على الخط المباشر عملية مكلفة وتتطلب عملاً مضنياً. وتستطيع المكتبات توفير أعبائها من خلال تضافر جهودها في إتاحة هذه المجموعات في شكل اتحادات. وسيؤدي ذلك إلى توفير جهود الناشرين؛ حيث سيكون لديهم عددٌ أقل من العملاء الذين يتفاوضون معهم ويدعمونهم. وفي الولايات المتحدة تقوم الولايات (كولاية أوهايو على سبيل المثال)، والشركات الأكاديمية بتنظيم اتحادات المكتبات. وفي أوربا، حيث تدار معظم الجامعات عن طريق الدولة، هناك عدد من الاتحادات الوطنية الشهيرة التي تعمل على تقديم خدمات المكتبات الرقمية للمجتمعات الأكاديمية كافة. ولعل مشروع ميلفيل Melvyl، هو أحد الأمثلة الجيدة على الجهود التعاونية الذي أقامته جامعة كاليفورنيا قبل انتشار الويب والذي يهدف أساساً إلى دعم تقاسم المكتبات لسجلات الفهارس وسجلات الكشافات).

اللوحة رقم (٥–٦)

مشروع ميلفيل ومكتبة كاليفورنيا الرقمية

(Melvyl and the California Digital Library)

تعمل التجمعات الأكاديمية التسعة التابعة لجامعة كاليفورنيا وكأنها تسع جامعات مستقلة، فالتجمعات الجامعية في بيركلي ولوس أنجليس تعد من الجامعات الكبيرة المستقلة، غير أنها من الناحية التنظيمية تمثل جزءاً من جامعة ضخمة. ولعدة سنوات تعاونت هذه التجمعات في إنشاء نظام مكتبة رقمية عُرفت باسم ميلفيل، وقد ظل كليفورد لينش Clifford Lynch يدير هذا

النظام إدارة صارمة في أطول فترات عمره.

ويتوافر في مركز نظام ميلفيل فهرس آلي يشتمل على جميع مقتنيات مكتبات التجمعات الأكاديمية التسعة، بالإضافة إلى مكتبة ولاية كاليفورنيا في سكرامينو، ومكتبة أكاديمية كاليفورنيا للعلوم. ويضم هذا الفهرس سجلات لأكثر من ٨,٥ ملايين عنوان، تمثل ١٣مليون مجلد. وبالإضافة إلى سجلات الكتب، يشتمل على سجلات ببليوجرافية لخرائط وأفلام ونوت موسيقية وملفات بيانات وتسجيلات صوتية. كما تضم قاعدة بيانات الدوريات حوالي ٨٠٠,٠٠٠ عنوان من المجلات والجرائد وأعمال المؤتمرات وغيرها، هذا فضلاً عن سجلات لمقتنيات المكتبات الكبرى الأخرى في كاليفورنيا. ويتيح نظام ميلفيل إمكانية الاتصال بالعديد من خدمات التكشيف والاستخلاص، بما في ذلك قاعدتا مدلاين Medlline وإنسبك Inspec المعروفتان. وفي عام ١٩٩٥م أضاف النظام صوراً من صور خرائط البتات bit-mapped images لمطبوعات معهد مهندسي الإلكترونيات والكهرباء IEEE، وربط هذه الصور بقاعدة بيانات إنسبك. وعندما يدخل المستفيد إلى قاعدة البيانات هذه تظهر له رسالة تقول : "الصورة متوافرة image available" حيث تشير إلى أن السجلات مربوطة بصور مطبوعات معهد مهندس الإلكترونيات والكهرباء IEEE المشار إليه سابقاً. وعليه يمكن للمستفيد أن يطلب من الخادم فتح نافذة على محطة العمل workstation الخاصة به لعرض تلك الصور. وكذلك يتاح استخدام الكتب والدوريات مجاناً للجميع، في حين يقتصر استخدام قاعدة مقالات الدوريات على منسوبي جامعة كاليفورنيا.

ويعد نظام ميلفيل من أول النظم التي تبنت تقنية المكتبة الرقمية الجديدة، وقد ارتبط تطوير المواصفة القياسية Z39.50 بهذا النظام. كما يعد فريق العمل بنظام ميلفيل مسؤولاً عن إنشاء شبكة الاتصالات التي تربط بين التجمعات الأكاديمية لجامعة كاليفورنيا.

لقد أثر نجاح نظام ميلفيل على تنظيم جامعة كاليفورنيا لخدمات مكتبتها الرقمية؛ حيث كان لكل تجمع أكاديمي من التجمعات التسعة مكتبته الخاصة به. وفي عام ١٩٩٧م وبعد عامين من التخطيط، قامت الجامعة بإنشاء مكتبة كاليفورنيا الرقمية التي تقدم خدماتها الرقمية للجامعة كلها. وقد كانت النية أن تكون هذه المكتبة من الناحية التنظيمية موازية للمكتبات الأخرى، وكان لمديرها ريتشارد ليوسر Richard Lucier ما لمدراء مكتبات التجمعات التسعة الأخرى من صلاحيات، وكانت ميزانيتها المبدئية حوالي ١٠ ملايين دولار في السنة مع توقع زيادتها بشكل كبير.

ويمكن للجامعة أن تعزز خدمات المكتبة الرقمية المركزية بسهولة، وذلك من خلال مناقشة اقتصاديات الحجم scale والتركيز على مفاوضات الترخيص والقوة الشرائية المدفوعة. لهذه الأسباب قامت التجمعات حسب تسلسلها التاريخي بتقاسم خدمات المكتبة عبر نظام ميلفيل الذي انضم إلى المكتبة الرقمية الجديدة. ولمكتبة كاليفورنيا الرقمية طموحات أخرى، وهي - بصراحة - مكلفة حيث تُكون جزءاً فاعلاً في التغيرات التي ستحول بنية نظام الاتصال العلمي، ويتوقع أن يكون لديها برنامج بحثي دقيق، وأن تعمل مع جهات أخرى في كل مكان من أجل تعزيز دور المكتبات في دعم نشاط

التعليم والبحث. وعلى أقل تقدير سوف تحصل جامعة كاليفورنيا على خدمات مكتبة رقمية ممتازة، وفي أحسن الظروف سوف تغير مكتبة كاليفورنيا الرقمية نمط الحياة الأكاديمية.

خدمات المعلومات الثانوية ومتعهدوها أو القائمون على تجميعها : aggregators

يتسع مصطلح المعلومات الثانوية secondary information services ليشمل مجالاً كبيراً من الخدمات التي تتمثل في إعداد الفهارس والكشافات والمستخلصات التي تعمل جميعها على مساعدة المستفيدين في الحصول على المعلومات. وقد نشأت بعض هذه الخدمات (مثل المستخلصات الكيميائية Chemical Abstract) في كنف بعض الجهات المهنية، في حين نشأ بعضها الآخر في كنف بعض الجهات التجارية مثل معهد المعلومات العلمي (International Scientific Institute (ISI)- الذي ينشر كلاً من خدمتي "المحتويات الجارية current contents "، و" كشاف الاستشهادات المرجعية science citation index "، والناشر بوكر Bowker المعروف -الذي ينشر دليل الكتب المتاحة في السوق Books in print)، أما مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC فله وضع خاص؛ حيث يعد منظمة شراكة membership organization تنمو حول عملية الفهرسة التعاونية.

إن هذه المنظمات كثيراً ما تكون عرضة للتغير وفي وضع يمكّنها من توسعة خدماتها لتصبح مكتبات رقمية، ومن خصائصها أن لها خبرة طويلة في مجال الحاسبات الآلية، والتسويق الجيد، وتوافر مجموعات ضخمة من البيانات القيمة. كما يتوافر للعديد منها الاحتياطي المالي القوي، أو أنها قد

تكون شركات تابعة لتكتلات تمتلك الأموال القادرة على دعم المشروعات الجديدة. ومن الملاحظ أن كل واحدة من هذه المؤسسات ترى بأن مستقبلها مرهون بالتكامل بين تقديم المعلومات الثانوية والأولية، ولذلك تنشأ كثير من المشروعات المشتركة بين الناشرين ومقدمي خدمات المعلومات الثانوية هذه.

أما المتعهدون أو القائمون على التجميع aggregators فهم جهات تقوم بتجميع المطبوعات من العديد من الناشرين وتقدمها للمستفيدين جملة واحدة، وعادة ما يكون التوزيع من خلال بيعها للمكتبات. ولبعض مؤسسات التجميع هذه جذورها في نظم الاسترجاع على الخط المباشر المبكرة، مثل دايالوج DIALOG، ونظام الاسترجاع الببليوجرافي المعروف اختصاراً بـ "بي آر إس BRS". حيث تقوم هذه الخدمات بالحصول على الترخيص باستخدام الكشافات وقواعد البيانات الأخرى، ثم تحميلها على حاسبات مركزية ذات واجهة بحث متخصصة تمتلك إمكانيات وصول قوية. وفي هذه الأيام يمكن لإحدى خدمات تجميع المعلومات الكبيرة أن تتفاوض بشأن الحصول على تراخيص من ٥٠٠٠ ناشر أو أكثر ثم تقدم ترخيصاً واحداً للعميل.

الجامعات ومكتباتها :

آفاق التغيير في المكتبات الجامعية:

تواجه الجامعات كغيرها من المؤسسات الأخرى صعوبة في التعامل مع التغيير؛ فالجامعات مؤسسات كبيرة ومحافظة وفيها من أسباب التباين الواضحة التي تعوق عمل الفريق. وإذا كان من غير الممكن تجاهل الفارق الثقافي بين الإنسانيات والعلوم، فإن مثل هذه الفوارق تضع حدوداً بين الباحثين والمكتبين واختصاصيي الحاسبات. يضاف إلى ذلك أن أعضاء هيئة

التدريس كثيراً ما يعاملون زملاءهم من غير أعضاء هيئة التدريس بازدراء [١]، كذلك للمكتبين لغتهم الاصطلاحية الخاصة، هذا في الوقت الذي يرى فيه اختصاصيو الحاسبات أن المعرفة الفنية هي المقياس الوحيد للتألق.

إن الأقسام الأكاديمية التي يهيمن على إدارتها أعضاء هيئة التدريس ممن ليس لهم هذا الحق tenured faculty memebers يشكل قوى جبارة تسعى للمحافظة على الأوضاع الراهنة، وإن إغلاق قسم معين، مهما كان ضعيفاً، ينظر إليه على أنه تخريب أكاديمي متعمد. ومن اللافت للنظر أنه عندما تغلق شركة مصنعاً لها، فإن أسعار أسهمها ترتفع، لكن عندما أغلقت جامعة كولومبيا كلية المكتبات بها هبطت أسهمها، مما يدل على ضعف الحوافز نحو التغيير، كما أن هناك معارضة قوية لهذا التغيير.

وحتى الآن لا تزال الجامعات مصدر الإبداع المستمر، وقد قدمت في الفصلين الثاني والثالث نماذج متعددة لجامعات تبنت تطوير تقنيات جديدة، وقامت على استخدامها واستثمارها قبل القطاع التجاري.كما أن تدفق شركات التقنيات العالية التي أنعشت الاقتصاد الأمريكي كان نتاج جهود عدد قليل من الجامعات البحثية تشمل جامعات: ستانفورد، وبيركلي، وهارفرد، وإم آي تي، وكورنيل، وكارنيجي ميلون.

إن الإبداع في إحدى المؤسسات الكبرى يتطلب إعادة توزيع استراتيجي للموارد؛ لأن المبادرات الجديدة تحتاج إلى اعتمادات مالية جديدة، ولا يمكن إيجاد هذه الموارد إلا من خلال اتخاذ خيارات صعبة. ومن الملاحظ أن

(١) لعل في هذا الكلام شيئًا من غير الموضوعية (المترجمان) .

الطريقة التي يتم بها إعادة توزيع الموارد في إحدى الجامعات البحثية تبدو عقيمة لأن تحويل الموارد من إحدى المناطق لبناء منطقة أخرى تسبب كثيراً من المشكلات. وقد قال مدير متحف أشموليان Ashmolean في أكسفورد ذات مرة ساخراً : " إن الاستراتيجية المثلى للمتحف هي أن تبيع جزءاً من مقتنياته حتى تتمكن من توفير التمويل اللازم للمحافظة على الجزء الباقي "؛ لكنه أردف قائلاً " إن المدير الذي سيقدم على تبني مثل هذه الاستراتيجية سيفقد منصبه ". وفي هذا السياق نجد قليلاً من عمداء الكليات من يؤيدون تقليص عدد أعضاء هيئة التدريس لتوفير الموارد التي من شأنها أن تدفع أعضاء هيئة التدريس الباقين للمزيد من الإنتاجية، ومع ذلك تتم الآن عملية إعادة توزيع الميزانيات. وسنة بعد سنة تزداد نسبة المخصصات المالية التي تنفق على الحاسبات والشبكات وموظفي الدعم، وربما تصل إلى ١% في كل مرة.

وليس من الواضح الآن، وعلى المدى البعيد، إن كان إعادة توزيع المخصصات المالية سيجلب مزيداً من الموارد للمؤسسات المكتبية الراهنة، أو ما إذا كانت الجامعات ستسعى إلى تطوير خدمات معلومات جديدة خارج مكتباتها، فالمؤشرات متضاربة، ومن المفارقة أن المعلومات الجيدة لم تكن تحظى بالأهمية الكبرى التي هي عليها الآن، ومع ذلك نجد أن مكتبات الجامعات تشهد تدنياً من حيث أهميتها مقارنة بمصادر المعلومات الأخرى. إن مكتبة الجامعة، بكل ما فيها من مجموعات الكتب والدوريات، ليست سوى جزء واحد في منظومة تبادل المعلومات الأكاديمية؛ فمنذ خمسين عاماً لم يكن يتوافر لأعضاء هيئة التدريس والطلاب بالجامعات إلا القليل من

مصادر المعلومات. أما اليوم فهناك العشرات من طرق الاتصال التي تتوافر لهم، فهناك الحاسبات الشخصية والإنترنت ومؤتمرات الفيديو المرئية Video Conferencing التي تسمح للباحثين بتبادل كميات كبيرة من المعلومات. وقد أصبحت التقنية بالبساطة التي يكون بمقدور الباحثين من خلالها إنتاج المعلومات وتوزيعها بمساعدة قليلة من المهنيين، كما يمكنهم القيام بذلك خارج حدود البنية الرسمية للمكتبة.

وإذا ما أردنا للمكتبات أن تكون المركز لخدمات المعلومات الجديدة، فلا بد لها من إعادة تخصيص مواردها داخلياً، لأن مناقشات ميزانيات المكتبات عادة ما تركز على التكاليف المرتفعة للمواد، وعلى الازدحام الشديد داخل المباني، وعلى تكاليف الحاسبات، مع أن أكبر التكلفة هي التي تنفق على العاملين. وقليلة هي الجامعات التي قامت بجهود صادقة لتقدير تكاليف مكتباتها، مع أن التقدير الحقيقي لإحدى المكتبات البحثية يجب أن يخصص حوالي ٢٥% من التكلفة للمقتنيات، وحوالي ٥٠% للعاملين، والنسبة المتبقية للمبنى والتجهيزات. وإذا أردنا للمكتبات أن تستجيب للفرص التي أوجدتها المعلومات الإلكترونية برغم ارتفاع الرواتب التي تجتذب العاملين المهرة، عليها أن تعيد أساليب إدارة وتنظيم العاملين فيها. فليس المطلوب هو فقط حث الناس على بذل المزيد من الجهد في أداء عملهم أو في تيسير الأعمال الداخلية، وإنما المطلوب هو ممارسة عملية إعادة بناء جوهرية.

مباني المكتبات الرقمية :

من مجالات التغير التي تنطوي على قدر كبير من الصعوبة للمكتبات بشكل عام، ولمكتبات الجامعات بشكل خاص، تلك المتصلة بكيفية التخطيط

لمباني المكتبات. ففي الوقت الذي تعد فيه المكتبات الرقمية محور التركيز في البحث والتطوير حول العالم، فإن أكبر مشكلة بالنسبة للعديد من المكتبات تتمثل في العجز الدائم في المكان. ففي ديسمبر ١٩٩٣م - على سبيل المثال - أصدرت مجالس تمويل التعليم العالي في المملكة المتحدة the funding councils for higher education in the UK، تقريراً حول مكتبات الجامعات، عرف بتقرير فوليت Follett Report، وكانت أهم توصية من وجهة النظر المالية هي الحاجة إلى برنامج لمبنى ضخم، وهي حاجة ملحة بصفة خاصة في بريطانيا، لأن عدد الطلاب في الجامعات ازداد بشكل كبير، وبالتالي زادت الحاجة إلى إيجاد أماكن للدراسة في التجمعات الجامعية.

وإن من الصعوبة بمكان تبرير الإنفاق على إنشاء مبنى جديد من أجل تخزين أوراق يندر استخدامها، غير أن العادات القديمة هي بطيئة التغير؛ فرغم إنشاء مكتبات في أماكن كثيرة، كالمكتبات الوطنية الجديدة في كل من لندن وباريس، لا تزال عادة الإبقاء على وجود فهرس بطاقي موضوع في أدراج من خشب البلوط الفاخر في كثير منها، من أجل تلبية حاجات كبار المستفيدين الجدد، وذلك رغم إتاحة مثل هذه الفهارس على الخط المباشر وعدم إضافة أية بطاقات جديدة إليها.

وللإفادة من المكتبة التقليدية، على المستفيد أن يذهب إليها، حتى إن كانت بعض المكتبات تسعى إلى توصيل الكتب أو صور من غيرها من مصادر المعلومات إلى بعض المستفيدين ذوي الاحتياجات الخاصة، فإن هؤلاء المستفيدين يجب أن يكونوا قريبين من المكتبة ومعروفين لدى

موظفيها. وعلى العكس من ذلك، فإن المستفيدين من المكتبة الرقمية ليسوا مضطرين للذهاب لمكان وجود المكتبة، لأن أمناء المكتبات ومديري مواقع الويب وغيرهم من المهنيين المسؤولين عن إدارة المجموعات يتواجدون في مكاتبهم ويؤدون عملهم من هناك، وليس هناك سبب يدعوهم لمقابلة المستفيدين. وعليه يمكن القول إن مكتبة نيويورك العامة لابد أن تكون في مدينة نيويورك، أما مكتبة نيويورك الرقمية فيمكن أن تختزن مجموعاتها في برمودا Bermuda.

إن من الصعوبة التكهن بما سيكون عليه المبنى الجيد للمكتبة في السنوات المقبلة، ولعل المحاولات والمحن التي عاشها مبنى المكتبة البريطانية الجديدة في لندن يظهر ما يمكن أن يحدث بدون التخطيط الجيد. وإذا كان من المعروف أن مبنى المكتبة النموذجي يعمر خمسين سنة على الأقل، فإن أحداً لا يستطيع التكهن بما ستكون عليه المكتبة الأكاديمية بعد سنوات قلائل من الآن. ولذلك فإن التركيز في تصميم مباني المكتبات الجديدة يجب أن ينصرف نحو فكرة المرونة. ومع أنه على مصممي مباني المكتبات أن يضعوا في الاعتبار متطلبات الاتصالات التي لم تتضح معالمها بشكل كامل حتى اليوم، فإن جميع الأماكن يجب أن تكون مزودة بكابلات توصيل الشبكات عامة الأغراض، وبالتمديدات الكهربائية الكافية، ورغم كل ذلك يجب أن تكون المنشآت نفسها ملائمة للمخازن التقليدية.

اللوحة رقم (٥ - ٧)

تحديث مكتبة هارفرد القانونية

The Renovation of the Harvard law Library

تعد مكتبة لانجديل هول Langdell Hall، المكتبة الرئيسية لكلية الحقوق في جامعة هارفرد، وهي ليست مجرد مكتبة لأكبر كلية حقوق، لكنها تعد أيضاً واحدة من أكبر المجموعات المتخصصة في الدراسات القانونية وتاريخها. وخلال عامي ١٩٩٦و ١٩٩٧م، تم تحديث المكتبة تحديثاً كاملاً بكلفة بلغت ٣٥مليون دولار. ويكشف هذا المشروع التحديات التي تواجه المباني على المدى البعيد في حقبة تتسم بالتغير المتسارع.

وأثناء عملية التخطيط لهذا التحديث قدمت جامعة هارفرد خطتين مقترحتين كبيرتين، تتمثل الأولى في عدم توفير أماكن بالمكتبة لحاسبات عامة، وتتمثل الثانية في توفير مقومات تشغيل حاسبات محمولة في كل مكان، ذلك على افتراض أن زوار المكتبة سوف يأتون بحاسباتهم الشخصية [المحمولة]معهم عند استخدامهم للمكتبة؛ لذلك تم تجهيز كل موقع للزوار بمنفذ للتيار الكهربائي وبمنفذ للوصل بالشبكة. وبموجب ذلك بلغ عدد منافذ الشبكة Ethernet Port ٥٤٠ منفذاً متوافراً على الطاولات وفي المقصورات والردهات وقاعات الدراسة في المبنى، كما تم إعداد ١١٥٨ وصلة شبكة بالمبنى وتم تشغيل معظم هذه الوصلات بشكل سريع (في حين تم وضع الجزء المتبقي في وضع الاحتياط). ولتخفيف قرار "عدم إتاحة حاسبات عامة" قامت الجامعة بتركيب حوالي مائة حاسب عام، بالإضافة إلى غرفة دراسية واحدة للحاسبات، ومعملين صغيرين لأغراض التدريب.

وتعد تكلفة هذا المشروع التي بلغت ٣٥ مليون دولار أمريكي مبلغاً كبيراً حتى بالنسبة لكلية هارفرد للحقوق. ولاشك أن الكلية قد غامرت على افتراض أن جميع المستفيدين يمتلكون حاسبات محمولة، وعلى أن

إعدادات الشبكة لديها من المرونة الكافية للتكيف مع أية تغيرات في المستقبل. وبالرغم من ذلك فإن أكبر مغامرة كانت تلك التي لم يتم الإفصاح عنها، وهي أن المكتبة سعت بالفعل إلى زيادة المساحات المخصصة لمستخدمي المكتبة. وكان الافتراض أن هؤلاء المستفيدين سوف يستمرون في القدوم للمكتبة لأغراض الدراسة مع أنه من المعروف أن أعضاء هيئة التدريس بكلية الحقوق يفضلون العمل في مكاتبهم أكثر من ذهابهم للمكتبة. وقد قال أمين المكتبة هاري إس مارتن الثالث Harry S, martin III منذ عدة سنوات مضت: "إن الهدف هو إبعاد أعضاء هيئة التدريس عن المكتبة"، ولعل هذه المقولة كانت تعهدًا بتقديم الخدمة لأعضاء هيئة التدريس في مكاتبهم، بما في ذلك خدمات توفير المعلومات المباشرة، وخدمات توصيل مصادر المعلومات courier service الجيدة، ومع ذلك فإن الاهتمام بأماكن القراءة في مكتبة لانجديل يعني أن هناك اعتقاداً بأن الباحثين القانونيين وطلاب كلية الحقوق سوف يأتون إلى المكتبة للقيام بعمل جاد لسنوات عديدة.

لم تذهب الأموال كلها إلى التقنية، فمكتبة لانجديل تقع في مبنى قديم رائع، وتضم كتباً تاريخية، وفنوناً قيمة، وإنما الجزء الأكبر من الميزانية ذهب في أعمال المصاعد والتدفئة والتكييف والسباكة. وفي هذا السياق لم تنس المكتبة مكانتها وعراقتها؛ حيث تم إصلاح المناضد القديمة، واستبدلت الشمعدانات بمصادر إضاءة هابطة غير مؤثرة، وأعيد ترتيب النقوش اللاتينية، وتم توضيح النقوش ضئيلة البروز bas-relief symbols، وإن هذه الرؤية لمكتبة قانونية عظيمة تجمع بين النظرة التقليدية للمكتبة

مع الاستفادة من التقنية الحديثة.

الفصل السادس

قضايا اقتصادية وقانونية

يعد هذا الفصل- أكثر من أي فصل آخر- استعراضًا سريعاً لبعض القضايا التي هي في غاية التعقيد، مع ملاحظة أننا لم نحط بكل تلك القضايا بل لم نحاول القيام بذلك، وما سيرد في ثنايا هذا الفصل يعكس وجهة نظري الشخصية، التي أرى أنها في حاجة لمراجعة مستمرة مع مرور الوقت، وإن كنت آمل أن تظل الأفكار الأساسية التي تحملها وجهة النظر هذه مطروحة كما هي.

إن المكتبات الرقمية تتطور في ظل عالم تشابكت فيه قضايا المستفيدين والمحتوى والتقنية مع الأبعاد الاقتصادية والاجتماعية والقانونية المحيطة بتلك القضايا، ومن ثم فإن دراسة أي من هذه القضايا لاينبغي أن تتم بمعزل عن القضايا الأخرى، بل ينبغي دراستها في سياق شامل ينظر إليها جميعاً، فكل من صناعة النشر وعالم المكتبات تحكمهما قواعد عمل وأعراف قطعت سنين طوال حتى صارت على النحو الذي نراها عليه اليوم، ومع اتساع مجال النشر الإلكتروني وانتشار المكتبات الرقمية، تسارعت وتيرة التغيير في ممارساتها التجارية وفي الإطار القانوني الذي يحكمها.

وحيث إن التقنية هي القاعدة أو الركيزة الأساسية للمكتبات الرقمية، فإن البعض يعلق آماله على الحلول الفنية التقنية لتتكفل بحل جميع المشكلات، في حين يعتقد البعض الآخر بأن حل كل شيء يكمن في سن قوانين جديدة، وأرى أن كلا الرأيين غير موفق؛ فالتقنية يمكن أن تسهم في تقديم الحلول،

وليس بمقدورها أن تحل جميع القضايا القانونية والاجتماعية، كما أن تغيير القوانين قد يفيد لكن إيجاد القوانين الضعيفة أسوأ من عدم وجود القوانين؛ فالقوانين تكون فاعلة إذا كانت تصوغ إطاراً يفهمه الناس، ولديهم من الاستعداد ما يدفعهم لتقبل هذا الإطار، وبالمنطق نفسه، يمكن القول بأن النماذج التجارية للمعلومات الإلكترونية لن تجدي ما لم تتوافق معها مصالح جميع الأطراف ذات الصلة بالموضوع. وعليه فإن التحدي الحقيقي يتمثل في إيجاد أعراف اجتماعية للإفادة من المعلومات أو استخدامها شريطة أن تكون مفهومة ومقبولة من قبل الشريحة الكبرى من الناس بشكل عام، فإذا استطاعت هذه الأعراف أن تتيح لعدد معقول من الناس إمكانية أداء أعمالهم، وكان ذلك موضع ملاحظة الآخرين، فسيتلو ذلك الأطر الاقتصادية والقانونية.

القضايا الاقتصادية :

يتسم عالما المكتبات والنشر بأنهما من مجالات العمل ذات الطابع التجاري الكبير؛ فهناك مصانع ضخمة تنتج المعلومات ومواد التسلية من أجل الكسب المادي، مثل المواد الفيلمية للتحقيقات المصورة، والصحف وصور الإعلانات التجارية، والروايات، والكتب الدراسية، وبرمجيات الحاسب الآلي، والتسجيلات الموسيقية وغيرها. وتشير بعض التقديرات إلى أن هذه الصناعات تشكل ٥% من إجمالي اقتصاد الولايات المتحدة الأمريكية، ففي عام ١٩٧٧م وصلت ميزانية مكتبات جامعة هارفرد إلى أكثر من خمسين مليون دولار، وبلغ عدد من يعملون بتلك المكتبات إلى ما يقرب من ألف موظف. أما مكتبة الكونجرس فيبلغ عدد موظفيها نحو ٤٥٠٠ موظف،

وتمثل شركة "تايم وارنر Time Warner" وغيرها من دور النشر قوى كبرى في الأسواق العالمية للأسهم. وفي عام ١٩٩٦م دفعت شركة "تومسون Tomson corporation" ما يزيد على ثلاثة بلايين دولار لشركة "ويست ببلشنج West publishing"، كما تقدر أسهم شركة "ياهو Yahoo" المعروفة بأكثر من ذلك، رغم ضآلة أرباحها.

وتنبع اهتمامات الناشرين بوضع الإطار التجاري والقانوني للمعلومات المباشرة المتاحة على الإنترنت من رغبتهم في تحقيق مكاسب هائلة من ناحية، والخوف من ضياع تلك المكاسب من الناحية الأخرى، وتأتي تلك الرغبة من اعتقاد هؤلاء الناشرين أنه بإمكانهم تحقيق مبالغ ضخمة من الأموال من وراء المعلومات الإلكترونية إذا استطاعوا معرفة السبيل إلى ذلك. أما الخوف فمصدره أن صور التغيرات الاقتصادية ستقضي على مصادر الدخل التقليدية، وستؤدي إلى زوال المنظمات، وإلى فقدان الكثيرين لوظائفهم؛ وهو تخوف ـ على أية حال له ما يبرره تماماً لدى المختصين في مجال الحاسب الآلي؛ فمعظم الشركات التي حققت نجاحاً في أحد أجيال الحاسبات الآلية منيت بالفشل في الجيل الذي يليه، حيث إن شركات الحاسبات المركزية Mainframe، مثل يونيفاك Univac، وسي دي سي CDC ، وبورجس Burroughs، وهني ويل Honeywell، اختفت وحلت محلها شركات الحاسبات الصغيرة Minicomputer، كما أن الشركات التي تخصصت في هذا النوع من الحاسبات، مثل شركة "برايم prime، ووانج Wang، وداتا جنرال Data General" لم تستطع أن تقاوم الانتقال إلى صناعة الحاسبات الشخصية، كما لم تعمر طويلاً أوائل الشركات المتخصصة في

هذا النوع، كما هو الحال مع معظم رواد البرمجيات، بل إن شركة آي بي إم IBM فقدت سطوتها على السوق، وهنا نتساءل: هل ستشهد المعلومات الإلكترونية المصير نفسه ؟، فالناشرون ومتعهدو خدمات المعلومات ممن هيمنوا على الأسواق التقليدية قد لا تكون لهم الريادة بالضرورة في الأسواق الجديدة.

وأيًا ما كان نوع الشركة، كبيرة أم صغيرة، تجارية أم غير ربحية، فإن عليها أن تغطي تكاليفها، فكل مرحلة من مراحل إنشاء المكتبات الرقمية، وتوزيع خدماتها واستخدامها والإفادة منها، تكلف الكثير. فالكُتّاب والمصورون والقائمون على صف الحروف والمصممون والمحررون، كل أولئك يريدون مقابلاً عن جهودهم في إنتاج المعلومات. وفي كثير من الأحيان، وإن لم يكن جميعها، يكون هذا المقابل مالياً؛ فالناشرون وأمناء المكتبات والأرشيفيون وباعة الكتب ووكلاء الاشتراكات، والمتخصصون في الحاسب الآلي، جميعهم يطمحون إلى الكسب المادي، وإلى الآن ليس ثمة إجماع على أفضل طريقة لدفع قيمة المعلومات الإلكترونية المتاحة على الإنترنت، بل إن كل طريقة ممكنة يتم إخضاعها للتجريب تقريباً. ودون الدخول في تفاصيل، يمكن تقسيم تلك الطرق إلى فئتين: فئة ترتبط بالوصول المتاح مجاناً (حيث يتحمل منشئ أو منتج المعلومات التكلفة)، وفئة ترتبط بالوصول المقيد restricted access (حيث يدفع المستفيد أو مكتبته مقابلاً عن الوصول إلى المجموعات).

لقد أشرنا في الفصل الخامس إلى تنوع دوافع المهتمين بالمكتبات الرقمية والنشر الإلكتروني، ولاحظنا على وجه التحديد كيف أن منتج

المعلومة قد تكون له أهداف مالية تختلف عن تلك التي تكون لناشرها، وعندما تكون دوافع منتج المعلومة مالية بالدرجة الأولى فإنها تتفق تمام الاتفاق مع دوافع الناشر، وهي الحصول على العائد وكسب الربح، لكن السؤال هو: ما نصيب كل منهما من ذلك العائد؟ فإذا كان دافع منتج المعلومات غير مالي، بينما يركز الناشر على تغطية التكاليف، أو تحقيق الأرباح لأصحاب الأسهم، فقد تتعارض الأهداف فيما بينهما.

والجانب الجديد بالاعتبار بالنسبة للويب، أن هناك قدراً هائلاً من المواد الجيدة متاحة مجاناً دون أية تكاليف يتحملها المستفيد. ولقد سبق أن تنبأ الكثيرون بأن "الوصول المتاح مجاناً open access" قد يحد من جودة المواد المتاحة على الويب، بينما الواقع اليوم يقول غير ذلك؛ فهنالك قدر لا يستهان به من المعلومات عالية الجودة متوافرة على الشبكة ويتحمل تكلفتها ويشرف على صيانتها منتجوها. وباستعادة الأحداث الماضية فليس في ذلك ما يثير الدهشة، لأن كثيراً من منتجي المعلومات وموزعيها يرغبون في نشر ما لديهم من معلومات وإطلاع الآخرين عليها، كما أن لديهم استعداداً لأن يتحملوا تكلفة ذلك على حسابهم الخاص. ومن أمثلة من يستثمرون مواردهم الخاصة في إتاحة ما لديهم من مواد مجاناً دون قيود أولئك الباحثين الراغبين في كسب اعتراف الغير بتميزهم في مجالهم، وكذلك الهيئات الحكومية، وشركات الدعاية والإعلان، والهواة، وغيرهم ممن يشعرون في ذلك بالترويح عن النفس.

وكلما كانت الرغبة في نشر الأعمال والإفادة منها على نطاق واسع هي الدافع الأساسي لمنتجيها، كان الوصول المتاح مجاناً هو خيارهم المفضل، لكن

لابد من إيجاد المصدر الذي يعمل على تدبير الموارد المالية اللازمة للحفاظ على هذه المجموعة وصيانتها، وهنا تبرز "المنح" كأحد مصادر التمويل المهمة، حيث يعتمد مشروع بيرسيوس Perseus project بصورة كبيرة على المنح المقدمة من المؤسسات. وفي مكتبة الكونجرس تتولى إدارة المنح والصناديق المالية الداخلية تمويل مشروعي "الذاكرة الأمريكية American Memeory"، و "برنامج المكتبة الوطنية الرقمية The National Digital Library Program". صحيح أن المنح المقدمة عادة ما تكون قصيرة الأجل، إلا أنه يمكن تجديدها، كذلك يتلقى مشروع "أرشيفات لوس ألاموس للطباعة الإلكترونية los Alamos E-Print Archives" منحة سنوية من "المؤسسة الوطنية للعلوم National Science Foundation"، كما يتلقى مشروع "نت لب Netlib" منذ بدايته تمويلاً من "هيئة مشاريع الأبحاث المتقدمة للدفاع المعروفة بداربا DARBA"، والخلاصة أن التمويل عن طريق المنح لهذه المكتبات الرقمية يتخذ شكلاً مؤسسياً بمعنى أنه يتم عبر هيئات ومؤسسات.

وتوفر شركات البحث عن المعلومات على الإنترنت مثل : "ياهو Yahoo، وإنفوسك Infoseek، وليكوس Lycos، الوصول المتاح للجمهور، معتمدة في ميزانيتها على الإعلانات، وقد أعادت هذه الشركات استكشاف النموذج المالي الذي تستخدمه شركات الإعلانات التليفزيونية في الولايات المتحدة الأمريكية؛ فشبكات التليفزيون تتحمل تكاليف إعداد البرامج وتوزيعها مجاناً للمشاهدين الجالسين في بيوتهم دون أن يدفعوا ثمناً مباشراً لها، والتي تدفع تكلفة هذه البرامج هي الشركات المعلنة عن منتجاتها.

ومن النقاط الفاصلة في مسيرة المكتبات الرقمية ظهور برامج البحث

عن المعلومات على الإنترنت، وقد حاولت بعض شركات تقديم الخدمات أن تفرض رسوماً شهرية على تلك الخدمات لكن "مايكل مولدين Michael Mauldin" صاحب شركة ليكوس Lycos أصر على إتاحة الوصول المفتوح للجميع، وانشغل بالبحث عن مصادر ربح بديلة، وحيث أن شركته كانت تقدم تلك الخدمة مجاناً، فإن الشركات المنافسة لم تجرؤ على فرض أية رسوم نظير الخدمة نفسها. وهاهي اليوم خدمات البحث على الإنترنت مفتوحة أمام الجميع. وبعد أن عاش اقتصاد تلك الشركات بعض السنوات العجاف، بدأت تحقق أرباحاً من الإعلانات، ومن منح التراخيص لدعم الوصول المتاح مجاناً إلى خدماتها.

ويهتم الباحثون اهتماماً خاصاً بنشر بحوثهم على نطاق واسع، وتلجأ كثير من هيئات البحوث إلى الاستعانة بمواردها الخاصة لصيانة مواقع الويب، فهذه المواقع تنشر كثيراً من البحوث العلمية الأصيلة التي تنشر لأول مرة، أما سلاسل مسودات الإنترنت The Internet Draft series التي تولاها ائتلاف المعلومات الشبكية Coalition for Networked Information (CNRI) فيتم تمويلها بطريقة مختلفة؛ حيث تعطى رسوم حضور اجتماعات لجنة العمل الخاصة بهندسة الإنترنت Internet Engeeniring Task Force (IETF) كرواتب للقائمين على إدارة منشورات هذه اللجنة، ويتاح الوصول المجاني لكل هذه المعلومات.

وتعد الإدارات الحكومية مصدراً آخر على قدر كبير من الأهمية من مصادر المجموعات المتاحة مجاناً على الإنترنت؛ فهي تقدم كثيراً من المعلومات لفترات زمنية قصيرة الأمد (مثل بيانات تتبع الأعاصير التي تقدمها

هيئة الأرصاد الوطنية)، لكن كثيراً من مجموعاتها (مثل مجموعة المعاهدات الدولية التي يشرف عليها مركز الالتزام التجاري Trade Comliance Center بوزارة التجارة) لها قيمة طويلة الأمد.

وهناك جهات خاصة تقوم بإتاحة مجموعات مكتباتها الخاصة مجانًا، ومنها مجموعات رائعة مخصصة لبعض الأنشطة الرياضية والهوايات وأندية الهواة. وهناك من المجموعات الشعرية ما ينشرها أصحابها بشكل خاص، أو الروايات التي ينشرها مؤلفوها على الإنترنت، وقيام منتج المعلومة بتحمل تكاليف إنتاجها وإتاحتها ليس بالأمر الجديد. بل يمكن القول بأن ما سيكون عرضة للاستخفاف في نشر الكتب على اعتبار أنه نوع من الغرور، يمكن أن يكون في الغالب مصدراً للمجموعات الرائعة المتاحة على الإنترنت.

وعندما يرغب ناشر إحدى مجموعات المكتبات الرقمية في تحصيل عائد من المستفيدين، يقوم بتقييد عملية الوصول إلى تلك المجموعة، بحيث لا يتسنى للمستفيد الوصول إلى مواد تلك المجموعة إلا بعد أن يدفع المقابل المالي لذلك. ويقتصر تناولنا في هذا الفصل على الممارسات التجارية، أما الوسائل الفنية المتبعة لإدارة الوصول إلى المعلومات فسنتناولها في الفصل القادم.

لقد جرت العادة على أن تعتمد صناعة نشر الكتب والدوريات على ما يدفعه المستفيدون، وعند بيع نسخة من كتاب إلى مكتبة ما أو قارئ معين، فإن العائد يوزع على الموزع والناشر والمؤلف وغيرهم من المساهمين في إخراج الكتاب، وكذلك الحال مع الأفلام السينمائية؛ حيث تُغطى تكاليف

الإنتاج والتوزيع من قبل المشاهدين وذلك من خلال بيع التذاكر في العروض السينمائية واستئجار أشرطة الفيديو. وإذا ما انسحب ذلك على المعلومات المتاحة على الإنترنت نصل إلى القول بفرض رسوم على الإفادة منها، وهو ما يحدث مع معظم المستفيدين من خدمات الاستشارات القانونية التي يقدمها موقعا ليكسيس Lexis وويسلو Weslaw، حيث يقومون بدفع مبلغ مالي يحدد حسب عدد ساعات إفادتهم من تلك الخدمة. وهناك وسائل بديلة جُربت لوضع رسوم على المعلومات المتاحة على الإنترنت يتم بمقتضاها تحديد التكلفة حسب الإفادة في ساعات الذروة (والتي قد تحدد بعدد الحاسبات التي يمكنها الوصول للمعلومة أو بأقصى عدد للمستفيدين منها في آن واحد)، إلا أن هذه الوسائل في ظل وجود اتفاقيات أو بروتوكولات الويب والإنترنت تحمل في طياتها شيئاً من التكلف أو التصنع.

ومن البدائل الأخرى أن يتم فرض رسوم حسب طبيعة المحتوى المُرسل للمستفيد، إذ يتيح بعض الناشرين الوصول إلى نص إحدى المقالات مقابل رسم معين قد يصل إلى عشرة دولارات يتم تحصيلها عن طريق بطاقة الائتمان، إلا أن التعامل عن طريق هذه البطاقات لا يلقى ترحيباً من المستفيد ويزيد من تكلفة هذه العملية. من هنا برز الاهتمام بإجراء بحوث لإعداد نظم الدفع الآلي automatic payment نظير المعلومات الموصلة عبر الشبكات، وتهدف هذه النظم إلى إيجاد خدمة تسديد الرسوم على الإنترنت تسمح بتعاملات آمنة الجانب من ناحية، ومنخفضة التكاليف لعمليات التحويل من ناحية أخرى، والأمل من ذلك هو تشجيع المنظمات الصغيرة على تقديم خدمات عبر شبكة الإنترنت دون الدخول في تعقيدات خدمات التسديد

الخاصـة. وإذا مـا تـم التوصـل إلـى مثل هذه النظم فإنهـا ستساعد علـى إيجـاد مجلدات ضخمة من سجلات التعاملات الصغيرة جداً. وفي الوقت الذي تظهر فيه المواد المادية (مثل الكتب) في صورة وحدات ثابتة، فإنه يمكن ظهور سوق للوحدات الصغيرة من المعلومات الإلكترونية.

ويعد مفهوم نظم الدفع الآلـي فـي الـوقت الحاضـر تـصوراً خياليـاً فـي الأساس، فالشكل السائد لتسديد قيمة مواد المكتبة الرقمية هو الاشتراك فيها، أي تسديد مبالغ مالية مجدولة ومحددة مقابل الوصول إلـى مجموعة من المواد، أما الاستخدام المفتوح فمسموح به طالما تم مراعاة بعض الضوابط المعقولة. وقد وضعت مجلة "وول ستريت Wall Street Journal" نظام بيع جيد تبيع بمقتضاه اشتراكات فردية تمكن المستفيد من الدخول علـى إصدارتها المتاحة على الإنترنت. كما تقدم الكثير من كبريات دور النشر العلمـي للمكتبات اشتراكات فـي دورياتهـا الإلكترونية. كمـا أن بعـض الجمعيات مثل "جمعية حرفيـي استخدام الحاسبات Association for Computing Machinery" تبيع اشتراكات للمكتبات وللأعضاء الأفراد. وقد بدأت بعض المكتبات الرقمية - التي أشرنا إليها في فصول سابقة - نشاطها وهي تعتمد على المنح، ثم تطورت بشكل تدريجي حتى وصلت للاكتفاء الذاتي في مواردها باعتمادها على الاشتراكات، وقد سار على هذا النهج كل من مشروع جستور JSTOR، واتحـاد الجامعـات للـبحوث الـسياسية والاجتماعية Inter-university Consortium for Political and Social Research.

ومرة أخرى يقدم لنا التليفزيون هنا نموذجاً جديراً بالاهتمام؛ فقد كان

للتليفزيون الأمريكي تجربة مع بديلين للعائدات الربحية من الإعلانات، تمثل أولهما في الدفع مقابل المشاهدة، وهو ما يستدعي من المشاهدين أن يدفعوا لكل برنامج يشاهدوه على حدة، إلا أن هذا الأسلوب قد أوجد سوقاً له وضعه اللائق لكنه لم يكتب له الانتشار الواسع. وأما البديل الثاني فهو ما تستخدمه شركات الكابل Cable Companies في مطالبة المشاهدين بتسديد اشتراك شهري لقاء مجموعة من البرامج، وقد حقق هذا النموذج العملي الثاني نجاحاً باهراً.

ويبدو أن مستخدمي المكتبات الرقمية، كشأن مشاهدي التليفزيون، يرحبون بدفع رسوم معروفة ومنتظمة؛ فالتسديد عن طريق الاشتراكات له مزاياه بالنسبة لكل من الناشر والمستفيد. كما أن المكتبات والمشتركين يعرفون التكاليف سلفاً، ولذلك يستطيعون وضع ميزانيتهم على نحو دقيق، كما أن الناشرين يعرفون حجم العائدات المتوقعة، ويرون أن الاشتراكات بالنسبة لهم تعوض أحد عيوب نظام التسعير القائم على الإفادة useo based pricing، بمعنى أن المواد التي تلقى طلباً مكثفاً تحقق أرباحاً طائلة، في حين أن المواد المتخصصة التي يقل الطلب عليها ليست بهذا الوضع.

ويعد نموذج الدفع عن طريق الاشتراكات subscription form of payment مصدر جذب للمكتبات، لأنه يشجع على الاستخدام على نطاق واسع، فالمكتبات تطمح إلى رؤية مجموعاتها موضع استخدام مكثف من قبل القراء وبأدنى قدر من العوائق. وقد ذللت المكتبات الرقمية كثيراً من العوائق المتأصلة في واقع المكتبات التقليدية، وسيكون من المؤسف وضع عوائق جديدة من خلال التسعير القائم على الاستخدام.

وهناك عامل اقتصادي يجعل المعلومات الإلكترونية تختلف عن غيرها من أشكال النشر التقليدية، وهو أن التكلفة تكاد تكون واحدة في كلا النوعين، سواء وجد من يستفيد من تلك المواد أم لم يوجد، حيث يتكرر كثير من المهام المرتبطة بإنتاج المواد، بما في ذلك عمليات الجذب والاختبار والتحرير والتصميم وتنسيق المواد في صيغ معينة، سواء كان المنتج في صورة مطبوعة أو صورة إلكترونية. إلا أنه في حالة المعلومات الرقمية تتضاءل تكلفة التوزيع بمجرد أن يتم إعداد النسخة الأولى من المادة؛ أي أن التكلفة من وجهة النظر الاقتصادية تكاد تكون ثابتة تماماً، كما أن التكلفة الهامشية تكاد تقترب من الصفر. ونتيجة لذلك فما أن تحقق مبيعات المادة الرقمية مستوى يغطي تكاليف إنتاجها، حتى تصبح كل المبيعات التالية بمثابة أرباح صافية، وما لم يتم تحقيق هذا المستوى، فإن هذه المادة يكون محكومًا عليها بالخسارة.

وفي حالة المواد ذات الطابع المادي (مثل الكتب والأقراص الموسيقية المدمجة أو الصور) يكون الأمر طبيعياً أن تضع ثمناً لكل نسخة مادية ملموسة؛ حيث يكلف إنتاج كل نسخة منها مبلغاً من المال، ويستشعر المشتري أن هذه التكلفة تنعكس بصورة من الصور في سعرها. أما في حالة المعلومات الرقمية، فإن من يقتنيها لا يتسلم شيئاً ملموساً في يديه، وهو أحد الأسباب التي تجعل من التسعير القائم على الإفادة غير ملائم. ويمكن القول إن الاشتراكات تتناسب مع التكلفة الثابتة نظير الإفادة من المواد في مقابل التكلفة الثابتة لإنتاج تلك المواد.

ومن العوامل الاقتصادية الأخرى التي تفرق بين النشر الإلكتروني

والنشر التقليدي، أن التسعير في النشر الإلكتروني يجب أن يستوعب تكاليف التغيير، صحيح أن تكاليف النشر الإلكتروني والمكتبات الرقمية باهظة حالياً، إلا أنها ستكون أرخص من تكاليف النشر التقليدي على المدى البعيد. وينبغي على المنظمات التي تريد دخول هذا المجال أن تواصل أنشطتها التقليدية، كما عليها في الوقت ذاته أن تضع استثماراتها في التقنية الحديثة رغم ارتفاع تكاليف النظم الجديدة، وبرغم أن حجم استخدام التقنية الحديثة لايزال محدودا نسبياً.

إن كثيراً من الإصدارات الإلكترونية، هي إصدارات لمواد صدرت في شكل مطبوع، وعندما يحدد الناشرون سعراً لإصدارات إلكترونية، فإنهم يرغبون في معرفة مدى تأثير مبيعات الإصدارات الإلكترونية على مبيعات نظائرها من الإصدارات المطبوعة، فمثلاً عندما يتاح قاموس معين على الخط المباشر، فكيف سيؤثر ذلك على مبيعات هذا القاموس في متاجر بيع الكتب؟ وعندما تتاح إحدى المجلات على الخط المباشر، فكيف سيؤثر ذلك على عدد اشتراكات الأفراد فيها؟ وبعد مرور عشر سنوات على تجارب النشر الإلكتروني، لم تزل الدلائل في طور التشكيل، ومن الصعوبة بمكان إيجاد دليل ثابت يحتكم إليه عند نشر إصدار معين. أما على المستوى الاقتصادي العام فمن الواضح أن المعلومات الإلكترونية غدت مصدراً هاماً يعمل له حساب في ميزانيات الاقتناء بالمكتبات، والإصدارات الإلكترونية والمطبوعة غالباً ما تتنافس بشكل مباشر من أجل كسب الرهان المالي الوحيد، فإذا تيسر لأحد الناشرين أن تكون عائداته من المعلومات الإلكترونية أكبر، فمن الممكن أن يكون ذلك على حساب الإصدارات المطبوعة. ولعل هذا التحول المالي من المواد المطبوعة لصالح الإصدارات

الإلكترونية يعد أحد القوى الضاغطة على الناشرين كي يتحولوا إلى النشر الإلكتروني.

وثمة بعض الأمثلة المثيرة والتي تتجسد في خدمات المعلومات الثانوية مثل: المستخلصات الكيميائية Chemical Abstract، والمحتويات الجارية Current Contents، اللتين غيرتا محتوياتهما الأساسية إلى حد ما منذ أواخر الثمانينات من القرن العشرين، حيث كانت السيادة فيهما للمواد المطبوعة أما الآن فالغلبة فيهما للإصدارات الرقمية والمتنوعة، وقد برع الناشرون في تقديم بدائل كثيرة من خدماتهم لكل من المستفيدين والمكتبات مثل الأسطوانات المدمجة، والأشرطة الممغنطة، والخدمات المتاحة على الخط المباشر، والمجلدات المطبوعة. وعلى أي ناشر يتخلف عن السوق الإلكتروني أن يتوقع تراجعاً ثابتاً في إيراداته، حيث سيجتذب هذا السوق الميزانيات التي كانت تنفق من قبل على شراء الإصدارات المطبوعة.

ومن حسن حظ أمناء المكتبات والناشرين أنهم غير ملزمين بالدفع لواحدة من أكثر مكونات المكتبات الرقمية كلفة وهي الشبكات ذات الأغراض العامة للحاسبات الشخصية The general purpose networks of personnel computers التي تقوم المؤسسات بتركيبها في كل مكان، حيث تأتي الاستثمارات في الاتصالات البعيدة وتجهيزاتها أو المعدات التي تشكل صلب المكتبات الرقمية من ميزانيات أخرى.

وليست الاشتراكات والتسعير القائم على الاستخدام أو الإفادة بالسبيلين الوحيدين لتحقيق العائد من المستفيدين، فإذا نظرنا إلى شركات التلفزيون الكابلي التي تعتمد على نظام الاشتراكات وجدنا أنها تعيد توزيع البرامج

التـي تنـتجها الـشبكات التليفزيونية، وقد لقـي هذا الـنظام معارضـة شديدة من شـركات الـبث، وقـد قـادت التكـتلات الـسياسية (خاصـة تلـك التـي قادهـا تـيد تورنـر Ted Turner) إلـى الوضـع الحالـي الـذي تـستطيع مـن خلالـه شـركات التلفـزيون الكابلـي التـي تعـتمد علـى نظـام الاشـتراكات، إعـادة توزيـع أي بـرنامج، لكن عليها أن تدفع نسبة من عائداتها فـي صـورة حصة ثابتة كلما عرضـت هذا البـرنامج أو ذاك، وعـندما تذيـع إحـدى محطـات الـراديو مـثلاً موسيقى مسجلة، فإن توزيع إجمالي عائد مؤلفيها وعازفيها واستديو التسجيل وغيرهم ممن ساهموا في إخراج هذا العمل الموسيقي، يعتمد في حسابه على نظـام معقد مـن نظـم المعايـنة Sampling، كمـا يحـصل اتحاد الكـرة الإنـجليزي علـى عائد المراهنة عن طـريق اسـتغلاله لحقوق طـبع قائمـته المثبـتة للفرق المشاركة في الدوري.

ولـيس هناك شيء قاطع بصدد أي من هذه الأساليب، فهي جميعها بمثابة حلـول مبـرمجة لأسئلة معقدة تتمـثل فـي كيفية قيام المستفيدين من المعلومات بتعويض من ينتجونها ويوزعونها.

دراسة حالة: الدوريات العلمية في شكلها الإلكتروني :

تقدم الـدوريات العلمية في شكلها الإلكتروني دراسة حالة مثيرة، لأنها من المجـالات الـرائدة التـي تجنـي فيها الإصـدارات الإلكترونية عائداً من المكتبات ومـن المستفيدين. كمـا أنها تلقي الضوء على التوتر القائم بين كل من الناشرين التجاريـين (الـذين يكـون هـدفهم الأساسـي تحقـيق أكبـر عائـد مـن الأربـاح)، والمؤلفين (الذين يطمحون أن يروا أعمالهم موزعة بصورة أوسع).

صحيح أن الناشرين الأكاديميين والمكتبات الجامعية شركاء طبيعيون، غير أن الاتجاه نحو النشر الإلكتروني ينبش جرحاً دفيناً بينهما، وكما هو واضح من اللوحة رقم (٦-١) حيث تستطيع المكتبات أن تقول بأن الناشرين يطلبون مقابلاً ضخماً نظير ما ينشرونه من دوريات، رغم أن الجامعات - وفي حالات كثيرة - هي التي كانت سبباً في المشكلة. ويأمل كثير من الباحثين وأمناء المكتبات أن تقود المكتبات الرقمية إلى طرق جديدة للنشر العلمي تتيح للباحثين وصولاً أرحب للبحوث وبتكلفة أقل للمكتبات. وفي الوقت نفسه يتعرض الناشرون التجاريون لضغوط كبيرة من المساهمين معهم في رأس المال حتى يحققوا أرباحاً عالية عاماً بعد عام.

ويتصف النظام الاقتصادي الذي يحيط بالدوريات العلمية بالغرابة؛ إذ يتم تمويل الباحثين ومعظم المحكمين وكثير من التكاليف الأخرى في المكتبات من قبل دافعي الضرائب، ولكن تؤول حقوق النشر بعد ذلك للناشرين. ويجب أن يقع الجانب الأكبر من اللوم على الجامعات التي سمحت لهذا الوضع أن يتطور إلى هذا الحد. وطالما أن من يقومون بالبحث هم أعضاء هيئة التدريس بالجامعات، ومن يشتري الدوريات هي مكتباتهم، فإن مجرد إجراء تغييرات بسيطة في سياسات الجامعات تجاه هذا الوضع كفيلة بأن توفر ملايين الدولارات مع الإبقاء على تعويض مالي معقول للناشرين، وقد بدأت بعض الجامعات في الآونة الأخيرة بالعمل بعضها مع بعض لتصحيح هذا الوضع.

إن السبب الحقيقي وراء تقاعس الجامعات عن القيام بأي إجراء تجاه هذا الوضع يكمن في أن المقالات العلمية لا يتم نشرها ابتغاءً للتواصل في مجال

البحث العلمي فحسب، بل لزيادة شهرة مؤلفيها في تخصصاتهم. هذه الشهرة الأكاديمية عادة ما تتحقق من خلال نشر مقالات الدوريات والكتب المتخصصة. كما أن التميز المهني المبني على النشر العلمي عادة ما يترجم إلى مناصب وترقيات ومنح بحثية، وإن العقبة الكبرى في مشوار الحياة الأكاديمية هي أن يحقق لنفسه مكانة مرموقة، تعتمد أساساً في حالتنا هذه على الإصدارات العلمية المحكمة. صحيح أن هناك بعض أساتذة الجامعات يؤلفون كتباً دراسية تدر الربح الوفير في بعض الأحيان، لكن جميع الناجحين من أساتذة الجامعات يؤلفون البحوث العلمية (في مجال العلوم) والمنفردات (في مجال الإنسانيات). وهناك مزحة مشهورة تقول: "إن عميد كليتنا لا يستطيع أن يقرأ، لكن المؤكد أنه يستطيع أن يقوم بإحصاء الأرقام"، ويقصد بالذي يستطيع أن يحصيه العميد البحوث المنشورة في الدوريات العلمية المحكمة، ويسعد الناشرون أن يحققوا أرباحاً من خلال نشر أشياء يُحصيها العميد. كما يتوقع من كل عضو هيئة تدريس أن ينشر عدة بحوث كل عام سواء أكان فيها ما هو جدير بالنشر أم لا. وحيث إن مكانة الباحث تتحقق من خلال كتابته الكثير من البحوث، فهناك ما يدفع الباحث ويحفزه على تقديم عدة أوراق علمية تدور جميعها حول موضوع واحد حتى ولو اختلفت في محتواها، إذ يركز كل منها على جانب معين من هذا الموضوع. وقد أظهرت كثير من الدراسات أن معظم البحوث العلمية لا يتم الاستشهاد بها من قبل باحثين آخرين، بل إن كثيراً من البحوث العلمية تفتقر إلى أدنى درجات الأهمية.

ومع ذلك فهناك ما يشير إلى أن هذا الوضع في طريقه إلى التغيير؛ فالمكتبات الرقمية تعمل على توفير بدائل تتيح للباحثين اطلاع العالم على

أبحاثهم. صحيح أن عملية التحكيم لها قيمتها في استبعاد البحوث الضعيفة والكشف عن الأخطاء الواضحة، لكنها تبديد للوقت، حيث إن الإجراءات التقليدية من تحكيم وتحرير وطباعة وتوزيع تستغرق في كثير من الأحيان ما يزيد على العام الواحد. يضاف إلى ذلك أن المقالة المنشورة في مجلة علمية تحفظ على أرفف المكتبة وتكون متوفرة فقط للأشخاص المصرح لهم بالدخول للمكتبة والمتمكنين من عملية استرجاع المقال، أما البحوث المنشورة على الإنترنت فهي متاحة للجميع. وقد ظهر في بعض التخصصات نمط جديد من النشر، وهو نشر تقارير البحوث أو المسودات على الإنترنت وإتاحة الوصول إليها دون قيود، ثم إعادة نشر هذه المواد لاحقاً في الدوريات العلمية لأغراض الحفظ التاريخي أو الترقي في الحياة الوظيفية.

اللوحة رقم (٦ - ١)

اقتصاديات الدوريات العلمية

The Economics of Scientific Journals

لقد خضعت الأرباح التي يجنيها الناشرون التجاريون من وراء الدوريات العلمية لدرجة من التمحيص المتزايد في السنوات الأخيرة. ففي الولايات المتحدة الأمريكية تقوم الحكومة الفيدرالية باستغلال ما تجمعه من أموال الضرائب في تمويل البحث العلمي في الجامعات والمعامل الحكومية والمراكز الطبية وغيرها من مؤسسات البحوث العلمية. وتتمثل الطريقة التقليدية لتقرير نتائج مثل هذه البحوث في قيام الباحث بإعداد دراسة علمية

ثم تقدمها إلى إحدى الدوريات العلمية.

وتتمثل الخطوة الأولى في عملية النشر في قيام أحد المحررين، الذي قد يكون متطوعاً أو موظفاً لدى الناشر، بإرسال الدراسة لعلماء آخرين لتحكيمها، حيث يقوم المحكمون بدورهم بقراءتها- تطوعاً دون تلقي مقابل على ذلك - بهدف تحكيم دقتها وخلوها من الأخطاء، والتوصية بنشرها أو بعدم نشرها. وتعرف هذه العملية بعملية "التحكيم peer review"، وما أن يقرر المحرر نشر هذه الدراسة حتى يتيح لمؤلفها فرصة تعديلها وإعادة النظر فيها بناءً على مرئيات هؤلاء المحكمين.

ويضع معظم الناشرين قبل نشر الدراسات بعض القيود على المؤلفين منها، فعادة ما يُطلب منهم التنازل عن حقوق النشر لهم، والاشتراط عليهم كذلك بعدم إذاعة نتائج دراستهم قبل نشرها، وهكذا يتحكم الناشر في العمل دون أن يدفع شيئاً.

ورغم أن قليلاً من الدوريات (خاصة تلك التي تصدرها الجمعيات العلمية) لها اشتراكات من القراء، فإن سوقها الرائجة هي المكتبات الأكاديمية. كما أن كثيرًا من الدوريات ترفع قيم اشتراكاتها، إذ تصل في بعض الحالات إلى ما يزيد على ألف دولار سنوياً، وفي كثير من الأحيان ازدادت هذه الاشتراكات بنسبة تتراوح بين ١٠ - ١٥ % في التسعينات. ومع ذلك تجد المكتبات لزاماً عليها أن تشترك في تلك الدوريات لأن رواد هذه المكتبات من طلاب وأعضاء هيئة تدريس يبدون حاجتهم للوصول إلى البحوث المنشورة في تلك الدوريات.

الاشتراك في الدوريات الإلكترونية المباشرة Online Journals :

تدخل ضمن نطاق ناشري الدوريات العلمية في مجال العلوم والتقنية والطب شركات تجارية (مثل إلسفير Elsevier، وجون ويلي وأبناؤه John Wiley and Sons، وسبرنجر فيرلاج - Springer Verlag)، والجمعيات العلمية (مثل الجمعية الأمريكية لتقدم العلوم the American Association for the Advancement Science). وقد اندفع هؤلاء الناشرون بقوة إلى مجال النشر الإلكتروني، وأصبحوا من أوائل المنظمات التي تواجه التحديات الاقتصادية، ومع ذلك لا يزال توجههم ينصرف إلى نشر الإصدارات الإلكترونية جنباً إلى جنب مع نظيراتها التقليدية. ومنذ عام ١٩٩٦م أتاح كثير من الناشرين العلميين إصدارات إلكترونية على الإنترنت مناظرة لدورياتهم المطبوعة. صحيح أن الإصدارات الإلكترونية تتشابه مع الطبعات التقليدية لكنها ليست دائماً متطابقة تماماً، حيث يمكن أن في الإصدارات الإلكترونية بعض الأقسام الموجودة في الدوريات المطبوعة (مثل "رسائل إلى المحرر")، في الوقت نفسه يمكن أن تضيف بعض الإصدارات الإلكترونية معلومات إضافية لم يكن يتسنى إضافتها في الدوريات المطبوعة.

ويختار عدد من الناشرين نموذجاً مختلفاً عن النموذج الاقتصادي المألوف الذي يبيعون من خلاله اشتراكاتهم السنوية لمكتبات بعينها أو لاتحاد من المكتبات. ومن بين الناشرين الذين يتيحون للمستفيدين الاشتراك في كل من مطبوعاتهم التقليدية وإصداراتهم الإلكترونية على حد سواء، "المطبعة الأكاديمية Academic Press"، و "الجمعية الأمريكية للكيميائيين The American Chemical" و "اتحاد حرفيي استخدام الحاسبات The Association

For Computing Machinery " وإلسفير، و"الجمعية الأمريكية لعلماء الطبيعة The American Physical Society"، و "مطبعة جامعة جون هوبكنز The John Hopkins University Press" والناشر "سبرنجر – فيرلاج، ويلي Springer-Verlag and Wiley". أما مطبعة هاي واير High Wire فتقدم خدمات مباشرة لصغار الناشرين ممن تتميز دورياتهم المطبوعة بالجودة العالية، لكنهم لا يرغبون في استثمار أموالهم من خلال نظم الحاسب الآلي الباهظة الثمن. والنموذج الشائع هو أن يتيح الناشر إمكانية الوصول المجاني لفهرس محتويات الدوريات searchable index والمستخلصات لكنه يطلب مقابلاً مالياً إذا أراد المستفيد الوصول إلى النص الكامل للمقالة. ويمكن تحديد نظام الدفع إما عن طريق الاشتراك أو عن طريق الدفع مقابل كل مقالة fee per article.

وتتطلب الاشتراكات التي تتيح الوصول لمجموعة ما على الإنترنت اتفاقاً بين الناشر والمشترك، فإذا كان المشترك مكتبة يكون الاتفاق في صورة عقد مكتوب في العادة، ورغم اختلاف هذه الاتفاقيات من حيث التفاصيل، فإن هناك عدداً من القضايا تكون ثابتة في كل اتفاق، وسنسرد بعض القضايا في النقاط التالية :

– **المادة :** عندما تشترك مكتبة ما في دورية مطبوعة اشتراكاً سنوياً فإنها تتلقى نسخاً من الأعداد التي تصدر من تلك الدورية في تلك السنة، وقد تجد هذه النسخ طريقها إلى المخازن للأبد. أما عندما تشترك في دورية إلكترونية للسنة نفسها، فالمعتاد أن تحصل هذه المكتبة على تصريح بالوصول إلى هذه الدورية المتاحة على الإنترنت، وقد لا يقتصر الأمر على الوصول إلى أعداد السنة التي دفع الاشتراك عنها، بل قد يمتد إلى الأعداد السابقة التي

حوِّلت إلى صورة رقمية. وهكذا تكون المكتبة في حل من تخزين الأعداد السابقة، لكنها تفقد إمكانية الوصول إلى أي مادة ما لم تجدد الاشتراك في السنة التالية، أو إذا خرج الناشر من السوق.

– ***المستفيدون :*** إذا اشتركت إحدى المكتبات في مجموعة إلكترونية لأحد الناشرين عليها أن تحدد هوية المستفيدين من تلك المكتبة حتى يصرح لهم بإمكانية الوصول إلى هذه المجموعة، وبعض المكتبات لها مجموعة محددة من المستفيدين؛ لذا يكون من السهل نسبياً تحديد من له صلاحية استخدام المواد أو الإفادة منها سواء في المكتبة، أو في أحد فروعها، لكن هناك كثيراً من الجامعات والكليات بها أعداد كثيرة من الطلاب الذين يدرسون بشكل غير متفرغ، أو من أعضاء هيئة التدريس المرتبطين بالجامعة ولكن من خارجها عن طريق مستشفى تعليمي أو غير ذلك من المؤسسات المحلية التابعة للجامعة، كما أن هناك المكتبات العامة التي يجب أن تكون أبوابها مفتوحة للجميع كما يوحي بذلك اسمها. ومن حسن الحظ أن الغالبية العظمى من المكتبات لها إجراءاتها التقليدية في استصدار بطاقات عضويتها. ومن الطرق البسيطة في هذا الصدد أن تغطي اتفاقية الاشتراك كل من له حق الحصول على بطاقة عضوية المكتبة، وكل من هو متواجد داخل المكتبة.

– ***الأسعار المتاحة للمؤسسات متباينة الأحجام :*** هل من المعقول أن تدفع إحدى الجامعات التي يبلغ أعداد طلابها حول خمسة آلاف طالب قيمة الاشتراك نفسها التي تدفعها إحدى الكليات التي يبلغ أعداد طلابها نحو خمسمائة طالب، أو قيمة الاشتراك التي يجب أن تدفعها منظومة جامعات إحدى الولايات التي يصل عدد طلابها إلى مائة ألف طالب ؟ وما قيمة

الاشتراك التي يجب أن يتحملها معمل بحوث صغير في إحدى الشركات الكبيرة ؟ وهل يمكن لجامعة متخصصة في الآداب الحرة Liberal Arts أن تدفع قيمة الاشتراك في إحدى الدوريات الطبية التي لا تحتاجها إلا لفترات محدودة جداً كما تدفع القيمة نفسها إحدى كليات الطب التي تشترك في الدورية ذاتها؟ في واقع الأمر ليس هناك حتى الآن إجابات شافية لكل هذه الأسئلة.

– نظام التسعير المرتبط بنظم الاشتراكات في المواد المطبوعة :
عندما تكون المادة متاحة في صورة مطبوعة وفي صورة إلكترونية متاحة على الإنترنت في آن واحد، فكيف سيكون سعر كل منهما؟ أو كيف يمكن المقارنة بين السعرين؟، وعلى المدى البعيد ستصبح الإصدارات الإلكترونية أرخص من حيث تكاليف إنتاجها بسبب توفير تكاليف الطباعة والورق والتوزيع، لكنها على المدى القريب تمثل تحدياً كبيراً كمجال جديد من مجالات الاستثمار. وقد حاول قليل من الناشرين في بداية الأمر أن يفرضوا رسوماً أكبر على المواد الإلكترونية، أما الآن فهناك إجماع على فرض رسوم أقل إلى حد ما، ويحاول الناشرون أن يختبروا عدة خيارات من نظم التسعير لتشجيع المكتبات على الاشتراك في أكبر عدد من الدوريات.

– الطباعة : من أكثر النقاط إثارة للجدل وضع حدود لإفادة المشتركين من الدوريات المتاحة على الإنترنت. ومن أهم القضايا التي تثار في هذا السياق قضية طباعة تلك المواد، فناشر تلك الدوريات يجب عليه أن يتوقع قيام بعض المستفيدين بعمل نسخ مطبوعة من أعداد تلك الدوريات لاستخدامهم الشخصي، لكنهم محقون في اعتراضهم على قيام بعض المستفيدين ببيع ما

يطبعونه. ولم يتم الاتفاق حتى الآن على الأسلوب الذي يسمح بموجبه للأول طبع ما يريد، ويفرض فيه قيودًا على الثاني عند بيعه للمواد التي يصل إليها ويطبعها.

ومن الملاحظ أن نموذج اشتراكات المؤسسات أو الهيئات قد نقل عملية توصيل الدوريات العلمية من عالم الطباعة إلى الإنترنت دون معالجة أي من جوانب التوتر القائمة بين الناشرين وعملائهم، وسيكون من المثير حقاً أن نرى إلى أي مدى سينجح هذا النموذج في الصمود أمام اختبار الزمن.

الدوريات العلمية ومؤلفو مقالاتها :

يتجسد الفارق بين أهداف المؤلفين وأهداف الناشرين في القيود التي يضعها بعض الناشرين على المؤلفين؛ إذ يطلب كثير من الناشرين أن يضع المؤلفون قيوداً على توزيع مقالاتهم، ومن الطبيعي كذلك أن يحجم الناشرون عن نشر الأوراق العلمية التي نشرت في دوريات أخرى. بل إن كثيراً منهم قد يرفض نشر نتائج البحوث التي سبق أن أعلن عنها في أماكن أخرى كالمؤتمرات أو على مواقع الإنترنت، والبعض الآخر يسمح بوضع نسخة من البحث على أحد الحاسبات الخادمة كأرشيفات لوس ألاموس للطباعة الإلكترونيةLos Alamos E-Print Archives، لكن يطلبون حذف النسخة المتاحة مجاناً بمجرد نشر الطبعة المنقحة من المقالة. ولا شك أن هذا التصرف يثير ضجر المؤلفين، كما أنه ليس هناك ما يدل على انعكاس ذلك بالإيجاب على العائد من وراء النشر، بل يزداد هذا الضجر مع إصرار الناشرين على إلزام المؤلفين بالتنازل لهم عن حقوق النشر، مما يعني احتفاظ المؤلفين بأقل

الحقوق لجهود تأليفهم لهذه الأعمال، بل ويجرد دافعي الضرائب - الذين تمول أموالهم عملية التأليف والنشر برمتها- من أدنى الحقوق.

وقد حاولنا في "جمعية حرفيي استخدام الحاسبات (ACM)" أن نوجد توازناً بين مصلحة المؤلفين في نشر مؤلفاتهم على نطاق واسع والحاجة إلى تحقيق العائد المادي، وكان من نتيجة ذلك الوصول إلى سياسة أو مجرد محاولة مؤقتة في هذا الصدد، لكن يحدونا الأمل أن يكون هذا التوازن مقبولاً في السنوات القليلة القادمة. وقد أبقت الإصدارة الحالية من تلك السياسة على انتقال حقوق النشر كما هو المعتاد من المؤلف إلى الناشر، والتأكيد على إمكانية قيام الجمعية باستغلال المادة بأي هيئة وبأية طريقة تريدها، وفي الوقت نفسه تتيح للمؤلف مرونة كبيرة. وبشكل محدد، تشجع هذه السياسة المؤلفين على وضع مؤلفاتهم على حاسبات خادمة خاصة بهم سواء قبل نشر هذه المواد أو بعد نشرها. وفي المجالات سريعة التغير والتي تدخل في إطار اهتمام دوريات "جمعية حرفيي استخدام الحاسبات ACM" يتم في الغالب إتاحة المسودات- التي تكون دائماً على درجة كبيرة من الأهمية- مباشرة على الإنترنت مجاناً، بحيث يشرف على صيانتها إدارات الأبحاث أو أصحاب البحوث أنفسهم، ومع ذلك فلهذه السياسة مخاطرها؛ لأن المسودات تمثل أهمية لأصحابها، والجمعية تريد أن تشجعهم على ذلك، لكن هذا قد يفضي إلى تدمير سوق الدوريات العلمية على المدى البعيد.

الإطار القانوني :

صحيح أن هذا الكتاب ليس كتاباً دراسياً متخصصاً مجال القانون، رغم

خضوع هذا الجزء منه لمراجعة اثنين من المحامين، لكن ثمة قضايا قانونية تبلغ من الأهمية درجة تجعل عدم مناقشتها مدعاة للقصور في هذا الكتاب. وحيث إن النظام القانوني يوفر إطاراً يسمح بالتطوير المنظم للخدمات المباشرة المتاحة على الإنترنت، فليس هناك من غرابة في ارتباط جوانب قانونية كثيرة بالمكتبات الرقمية، وتشمل هذه الجوانب مجالات كثيرة، مثل: العقود وحقوق التأليف وغيرها من حقوق الملكية الفكرية، هذا فضلاً عن قضايا القذف والتشويه والمواد المنافية للآداب، وقانون الاتصالات، والخصوصية، والضرائب، والقانون الدولي.

ويتسم الوضع القانوني في الولايات المتحدة الأمريكية بالتعقيد بسبب تعدد جهات الاختصاص والتشريع مثل الدستور والمعاهدات الدولية، والقوانين الفيدرالية والحكومة، والحالات السابقة التي مرت على المحاكم في كل تلك المستويات. كما أن هناك كثيراً من القوانين التي يعمل بها في حدود إحدى الولايات كالقوانين التي تحظر نشر المواد المنافية للآداب، وهي لا تسري على ولايات أخرى، فقد حدث أن قام البعض بإتاحة مواد إباحية على بعض الحاسبات الخادمة في ولاية كاليفورنيا حيث يسمح قانون الولاية بذلك، إلا أنهم أدينوا على ذلك في ولاية لويزيانا التي تجرم مثل هذا العمل،كذلك هناك بعض القضايا القانونية الأخرى كحقوق التأليف التي تقع تحت تنظيم اللوائح الفيدرالية. لكن بعض أهم الموضوعات المتصلة بتلك القضايا لم تجد لها أي تفسيرات لدى المحكمة العليا الأمريكية، وعندما تصدر من هي دونها من المحاكم أحكاماً قانونية، لا يلزم ذلك غيرها من المحاكم، بل لا تجد غضاضة في تفسير القانون بشكل مختلف.

ويتميز كلا الفريقين اللذين يضطلعان بمسؤولية إنشاء المكتبات الرقمية على الإنترنت، وهما: علماء الحاسب الآلي، والمتخصصون في علم المعلومات، بأنهما يستخدمان الموارد المتاحة على نحو مشترك Shared Resources استخداماً مسؤولاً، إلا أن لكل منهما أعرافه وتقاليده في هذا الصدد، غير أن محاولة الدمج بينهما تخلق بعض المشكلات، لكن المشكلة الأكبر نتجت عن أن مجال الإنترنت قد اجتذب من ليس لهم مثل هذه الأعراف، بل إن هناك فريقاً من هؤلاء ذو نوايا سيئة، وفريق آخر لا يحسن استخدام مبدأ الحرية الموجود في عالم الإنترنت، بل هناك فريق ثالث لا يكترث مطلقاً بأن يكون سلوكه مسؤولاً ولا يشغله ذلك على الإطلاق.

وحتى أواخر الثمانينات كانت الإنترنت حكراً على الأكاديميين وغيرهم من الباحثين، وكانت هناك سياسات تنظم تحديد من له حق استخدامها، ومعايير الاستخدام الملائم لها. والأهم من ذلك كله أن المستفيدين اعتادوا مراقبة أنفسهم، وكان من ينتهك عرفاً معيناً يتعرض لسيل فوري من الاحتجاجات وأشكال اللوم الصريحة من زملائه، بالرغم من أن الأعراف الاجتماعية لم تكن رسمية إلى حد ما، إلا أنها كانت موضع احترام لأن معظم المستفيدين ينتمون إلى مجموعة صغيرة يتوارثون الأعراف فيما بينهم.

ومع انتشار الإنترنت بدأت هذه الأعراف تتهاوى، وقد رصدنا بالفعل بعض التغيرات في جامعة كارنيجي ميلون عندما التحق بها بعض الطلاب الذين كانوا قد تلقوا تعليمهم عن استخدام الحاسبات الشبكية Networked Computing بعيداً عن الجامعات وقبل التحاقهم بالجامعة نفسها.

وهناك كثير من القضايا القانونية لا تقتصر على المكتبات الرقمية، بل

هي عامة وتتصل باستخدامات الإنترنت، ويعكف المتخصصون في الإنترنت حالياً على إيجاد أساليب فنية للتحكم في رسائل البريد الإلكتروني التافهة التي تسبب عبئاً على طاقة الإنترنت وتعمل على إزعاج المستفيدين، كذلك تعد المواد الإباحية وعمليات المقامرة من المجالات التي حرضت على جني المكاسب المادية التجارية على حساب الأعراف الاجتماعية، كما أنها أيدت أهواء دعاة الحقوق المدنية ضد أخلاق الجماعات المحافظة.

تساؤلات دولية :

بحكم الانتشار الواسع للإنترنت، يمكن الوصول لمجموعة أي مكتبة رقمية من أي مكان في العالم، غير أن ما قد يعتبر تصرفاً طبيعياً للغاية في بعض البلدان، قد يعتبر مخالفاً للقانون في بلد آخر. فالولايات المتحدة على سبيل المثال تسمح بحيازة السلاح لكنها تحد من استخدام برامج التشفير Encryption Software في حين أن معظم الدول الأوربية لها موقف معاكس لهذا الموضوع.

كما أن الموقف من حرية التعبير يختلف اختلافاً كبيراً من بلد لآخر؛ ففي كل دولة بعض القوانين التي تحد من حرية التعبير والوصول إلى المعلومات، مثال ذلك القوانين التي تغطي مسائل السب والقذف، والمواد المنافية للآداب، والخصوصية والكراهية والعنصرية والأسرار الحكومية. وفي الولايات المتحدة ينظر إلى المعلومة على أنها حق ديمقراطي أصيل تكفله المادة الأولى من الدستور. كما أن المحاكم قد دأبت على التوسع في تفسير مفهوم حرية التعبير، ومع ذلك هناك حدود لا يمكن تجاوزها حتى في الولايات المتحدة الأمريكية، وكل قانون تم وضعه ينتظر منه أن يحكّم نشاطاً معيناً. وفي ألمانيا

هناك قوانين صارمة على مذهب النازية، كما أن في الدول العربية الصرامة نفسها مع كل ما يمس الذات الإلهية، ومع ذلك فمن الصعوبة بمكان التحكم في الإنترنت. ولسنوات طويلة كان هناك حاسب خادم في فنلندا يعمل على إرسال واستقبال الرسائل مجهولة المصدر على الإنترنت، لكن في النهاية وتحت الضغط الكبير من خارج البلاد، ألزمت المحاكم الفنلندية من يديرون هذا الخادم الإفصاح عن أسماء من يرسلون المواد الإباحية التي كانت ترد من جهة مجهولة من قبل.

إن الطابع الدولي الذي تكتسي به الإنترنت يخلق صعوبات في مسألة التجارة الإلكترونية بشكل عام بما في ذلك المعلومات الإلكترونية. وفي الولايات المتحدة تعمل القوانين على حماية كل من المستهلكين والموردين في تعاملاتهم التجارية التي تتم داخل حدود الولايات المتحدة، كما هو الحال في عمليات الدفع المالي، ومسائل البيع والشراء عن طريق البريد. غير أن الموقف يصبح أكثر تعقيداً عبر الإنترنت التي تغطي العالم بأسره، نظراً لأن موضوع التجارة هو المادة الرقمية التي يسهل نسخها أو تعديلها، يضاف إلى ذلك أن التعامل عبر الإنترنت لا يحتم على المتعاملين الإفصاح عن اسم الدولة التي تدار منها تلك العمليات.

المسؤولية القانونية :

المسؤولية عن محتوى المواد المكتبية من القضايا القانونية والاجتماعية ذات أهمية خاصة بالنسبة للمكتبات الرقمية، فالمجتمع ينتظر من مؤلفي المواد أن يتحملوا مسؤولياتهم إزاء مضمون ما يكتبون. كما أنه ينتظر ممن يحكمون على المضمون أن يتصرفوا بشكل مسؤول كذلك، لكن لن يكتب

للمكتبات الرقمية التطور والازدهار طالما بقيت المسؤولية القانونية معلقة في أعناق الأطراف التي تتركز وظيفتها الوحيدة في تخزين المعلومات وبثها.

وحيث إن للمكتبات مكانة مرموقة، فإن لها وضعاً قانونياً متميزاً في معظم الدول الديمقراطية، كما أنه من المستحيل، أو يكاد، أن نحمل المكتبة مسؤولية الآراء الخاطئة أو المواد الخارجة التي تتضمنها مجموعات هذه المكتبة. وهو المنطق نفسه الذي يسري على قانون الاتصالات بعيدة المدى الذي يحمي الشركات الناقلة للمعلومات، بحيث يحظر على أي منها مراقبة المحادثات الهاتفية التي تتم عبر خطوطها، بل إن شركات الهاتف محظور عليها التنصت عمداً على مثل هذه المحادثات.

وقد جرت العادة أن تقع المسؤولية عن محتوى المواد على المؤلفين، وعلى الناشرين الذين يجب عليهم أن يعرفوا قدر ما ينشرون، لا على المكتبات. وهو أمر أتاح للمكتبات أن تقتني المواد من كل الثقافات وعبر كل العصور دون الحاجة إلى تمحيص كل مادة خشية أن يكون فيها انتهاك لخصوصية البعض، أو حتى انتهاك لحقوق الملكية الفكرية للبعض الآخر [١]، وهناك من يرى أن هذه السياسة جيدة ويجب أن تنسحب على المكتبات الرقمية. وفي الوقت الذي تتحمل فيه المؤسسات مسؤولية المضمون الفكري للمواد التي تنتجها وتوزعها، ليس من المعقول أن تقوم المكتبات أو الشركات التي تقدم خدمات الإنترنت بمراقبة كل ما تبثه من مواد.

---

(١) للمكتبات بمختلف أنواعها سياسات ومعايير في اختيار مقتنياتها، وليس الأمر بهذه البساطة التي يصورها المؤلف (المترجمان).

إن المسؤولية التي تقع على عاتق الشركات التي تقدم خدمات الإنترنت من أهم النقاط التي يتناولها التشريع المنظم لحقوق التأليف في الألفية الرقمية الصادر عام ١٩٩٨م - والموضح في اللوحة رقم (٦-٢) - وهذا التشريع يعفي الشركات التي تقدم خدمات الإنترنت من انتهاك حقوق التأليف. كما يعفي كذلك المكتبات والمؤسسات التعليمية من ذلك. وكما هو الحال مع أي قانون جديد، فإن ظهور هذا التشريع لم يكن بالأمر الهين، كما أن المسألة لم تقف عند حد صدور التشريع، صحيح أن المكتبات الرقمية قد أبدت ترحيبها بهذا الجزء من التشريع، إلا أن كيفية تعامل المحاكم معه ستظل موضع متابعة ونظر.

اللوحة رقم (٦ – ٢)

تشريع حق المؤلف في الألفية الرقمية

The Digital Millennium Copyright Act

أقر الكونجرس الأمريكي عام ١٩٩٨م تشريع حق المؤلف للألفية الرقمية، وقد أدى هذا التشريع إلى تغيرات جوهرية في قانون حق المؤلف، وبصرف النظر عن الجزء الذي يتناول ما يتصل بتصميم بُدن السفن أو القوارب boat hulls، تكاد معظم مواد هذا التشريع تقتصر على الأعمال الرقمية المتاحة على الإنترنت. ويبدو أن التشريع يعمل على تحقيق توازن معقول بين المصالح التجارية التي يرغب أصحابها في بيع الأعمال الرقمية من جهة، وإتاحة المعلومات للجميع من جهة أخرى، وهو الأمر الذي يهم المكتبات ومرافق التعليم بشكل أساسي. وفي ظل هذا التشريع تتمتع الشركات التي تقدم خدمات الإنترنت في الولايات المتحدة بما في ذلك

المكتبات والمؤسسات التعليمية، بحماية كبيرة ضد الدعاوى القانونية بانتهاك حقوق التأليف، وهو أمر يحدث دون علم هذه الشركات، وحتى تتمتع هذه الشركات بهذه الميزة عليها أن تتبع قواعد محددة مثل تزويد المستفيدين بالمعلومات الخاصة بحقوق التأليف، كما أن عليها أن تتبع سياسة وقف الخدمة عن من تتكرر مخالفاتهم، فضلاً عن اتباع إجراءات معينة لحذف المواد المنافية للأخلاق واتباع الإجراءات الفنية المتعلقة بمعايير الصناعة؛ وأن يبلغ مكتب حقوق التأليف بوكيل معين بتسلم الإشعارات النظامية بمقتضى التشريع. كما يسمح هذا التشريع بشكل واضح بكثير من الأنشطة والأعمال التي هي من صميم أعمال المكتبات الرقمية، كما يسمح للشركات التي تقدم خدمات الإنترنت بتقديم تيسيرات للمستفيدين فيما يتصل بتخزين المواد كمواقع الويب، وبتتبع الروابط الفائقة، وباستخدام محركات البحث، ويقر التشريع كذلك بحق تلك الشركات في نسخ المواد للدواعي الفنية (كالحفظ لهذه المواد في حاسبات الذاكرة الخفية Caching، وبث هذه المواد أو توجيهها لمواقع أخرى). وفضلاً عن تلك الأعمال التي يقرها التشريع، فإنه يشير إلى عدم خضوع تلك الشركات للمساءلة عن انتهاكات المستفيدين إذا ما اتبعت القواعد بتصحيح المشكلات عند إشعارهم بها.

وبالنسبة للجامعات وغيرها من معاهد التعليم العالي يقدم التشريع استثناءً مهماً من قاعدة تحمل المؤسسات مسؤولية التصرفات التي يقوم بها موظفوها، ويتمثل هذا الاستثناء في عدم مراقبة أنشطة أعضاء هيئة التدريس بالجامعات وكذا طلاب الدراسات العليا المنتسبين لتلك الجامعات، وبالتالي عدم تحمل مسؤولية ذلك.

ويحظر التشريع تداول الأساليب الفنية التي يتبعها أصحاب حقوق التأليف لتقييد الوصول إلى أعمالهم، كما يحظر على القائمين بعمل هذه الأساليب وموزعيها محاولة كشف أسرار هذه الأساليب، إلا أن التشريع يسمح بعدة استثناءات تتميز جميعها بالتعقيد وتحتاج إلى تفسير دقيق، منها إمكانية قيام مطوري البرمجيات بتغيير نظم الحماية لكي تسمح بإمكانية التشغيل المتداخل، وإمكانية قيام الباحثين بدراسة نظم التشفير وأمن النظام. وإمكانية تفويض الجهات التنفيذية في نشر تقنية أمن الشبكة وإمكانية فحص المكتبات للمواد لتحديد إمكانية اقتنائها من عدمه، وأخيراً إمكانية السماح للمستفيدين بتجديد وإبطال مفعول الأساليب الفنية التي يمكن عن طريقها جمع المعلومات الخاصة عن المستفيدين وعن سلوكيات استخداماتهم.

ويقدم التشريع ضوابط للتأثير أو العبث بالمعلومات الإدارية الخاصة بحقوق التأليف لعمل معين، مثل عنوان هذا العمل أو مؤلفه أو ناقده أو مالك حق التأليف، فلا ينبغي تغيير هذه المعلومات أو إزالتها عمداً.

حقوق التأليف :

تعد العلاقة بين القضايا الاقتصادية وقانون حق المؤلف من أكثر القضايا القانونية التي يحتدم الجدل حولها عند الحديث عن المعلومات المباشرة على الإنترنت، ويبدو أن هذا الجدل دائماً ما يطفو على السطح كلما ظهرت تقنية حديثة. ومن الملاحظ أنه لم تكن هناك قوانين تتصل بحقوق التأليف عند اختراع الطباعة لأول مرة، وقد تعرضت مسرحيات شكسبير لعمليات قرصنة فكرية دون مساءلة. وفي القرن التاسع عشر قامت الولايات المتحدة بوضع قانون يتم بموجبه حماية حقوق المؤلفين الأمريكيين دون غيرهم. وقد

تعرضت أعمال كثير من الكتاب الأوربيين للنسخ في جرأة بالغة رغم صيحات الكثير من الكتاب من أمثال ديكنز Dickens وترولوب Trollope وغيرهم من الكتّاب.

وفي الولايات المتحدة كذلك تنسحب الحماية القانونية على كل المواد العلمية تقريباً، بما في ذلك المواد النصية والصور وبرامج الحاسب الآلي والمؤلفات الموسيقية والشرائط السمعية والبصرية، والاستثناء الوحيد من ذلك هو المواد التي يكتبها الموظفون الحكوميون، وفي البداية كان من ينتج العمل أو الجهة التي يعمل بها هي التي تمتلك حقوق التأليف، وبصفة عامة يمكن أن تعرض الملكية الفكرية للبيع والشراء شأنها في ذلك شأن أي شيء يُمتلك. وفي بعض البلاد هناك اتجاهات مختلفة إزاء الملكية الفكرية، ففي فرنسا وكندا تحديدًا، حقوق شخصية لصاحب العمل (يصطلح عليها بالحقوق الأدبية أو المعنوية) لا يمكن أن تورّث لشخص آخر.

وللحماية القانونية من الناحية الزمنية عمر محدد تنتهي فعاليتها بعد سنوات محددة من وفاة صاحب العمل. وقد قام الكونجرس الأمريكي بتحديد تلك الفترة عدة مرات، وكان آخرها عندما خشيت هوليوود من انتهاء فترة الحماية القانونية للأفلام القديمة، وهو مثال صارخ يجسد كم يأتي الصالح العام ثانوياً بعد المكاسب المادية لعدد قليل من الشركات.

ولمن يمتلك حق التأليف الخاص بأحد الأعمال الحق المطلق في نسخه واستخراج أعمال أخرى منه، وتوزيع نسخه سواء بالبيع أو بأي صورة أخرى، وهذا يسهم في طمأنة صاحب العمل من عدم المساس بعمله سواء

بقصد أو بدون قصد. كما يسمح للناشرين بتطوير منتجاتهم دون توجس من تأثر أسواقهم من وجود نسخ من مصدر آخر.

ورغم أن مالك حق التأليف يتمتع بسلطة كبيرة نحو استثمار العمل الذي قدمه، فإن قانون حق المؤلف ليس قانوناً مطلقاً، وهناك مفهومان هامان في القانون الأمريكي هما: البيع الأول First Sale، والاستخدام العادل Fair Use.

ويسري مفهوم "البيع الأول" على أي شيء مادي مثل "الكتاب"، بمعنى أن صاحب حق التأليف له السلطة في مسألة بيع كتاب جديد وفي تحديد سعره، إلا أنه بمجرد شراء أحد الزبائن نسخة من الكتاب انتقلت لذلك المشتري الملكية الكاملة للنسخة، لذلك من الممكن له أن يبيعه أو يتصرف فيه بأي شكل من الأشكال دون ضرورة الحصول على إذن بذلك.

أما مفهوم الاستخدام العادل فهو حق مشروع يسمح بالإفادة من المعلومات الخاصة بالعمل الذي يتمتع بالحماية دون إذن من صاحب الحق، وبناء على هذا المفهوم للاستخدام العادل يستطيع الباحث أو الدارس أن يقتبس فقرات قصيرة. كما يمكن له تصوير مقالة أو جزء من كتاب لأغراض الدراسة أو الاستخدام الشخصي. لكن حدود الاستخدام العادل غامضة بشكل متعمد رغم وجود أربعة عوامل رئيسية في هذا الصدد تتمثل في: الهدف من الاستخدام وطبيعة هذا الاستخدام (بمعنى هل يتم هذا الاستخدام لأغراض تجارية أم لأغراض تعليمية ليس المقصود منها الربح)، وطبيعة العمل الذي يتمتع بالحماية القانونية، وحجم الجزء المستخدم نسبة إلى حجم العمل الكلي وأهمية هذا الجزء، وأخيراً أثر الاستخدام على سوق العمل وعلى قيمته. ونظراً لافتقار هذه العوامل للدقة والتحديد، فإن القضاة يتمتعون بحرية تفسيرها.

وبصفة عامة يسمح الاستخدام العادل باستنساخ أجزاء من العمل، وليس العمل كله، وبتصوير نسخة واحدة لا نسخ كثيرة، وباستخدام شخصي لا استخدام تجاري، ولا يمكن إيضاح الفروق الدقيقة إلا في ضوء وجود سابقة قانونية. ومن الملاحظ أن إجراءات رفع الدعاوى القضائية أمام المحاكم تكلف مبالغ باهظة لدرجة أن كثيراً من قضايا حقوق التأليف الهامة (حتى فيما يخص المواد المطبوعة التقليدية) لم يتمكن أصحابها من إحالتها إلى المحاكم العليا، واكتفوا برفعها أمام محاكم صغرى.

وتجدر الإشارة إلى انه ليس من السهولة تطبيق مبدأ البيع أولاً، وحق الاستخدام العادل على المكتبات الرقمية. ففي الوقت الذي يمكن فيه تطبيق مبدأ البيع أولاً على الوسائط المادية التي تحمل عليها المواد الإلكترونية مثل الأسطوانات المدمجة، لا يمكن تطبيق ذلك على المعلومات المباشرة المتاحة على الإنترنت. كما أن الخطوط العريضة التي تحكم الاستخدام العادل يصعب نقلها من مجال الوسائط المادية إلى عالم الإنترنت. و قد كان هذا الغموض أحد الأمور التي أدت إلى سلسلة من محاولات إعادة صياغة قانون حق المؤلف، سواء في الولايات المتحدة أو خارجها. ورغم أن معظم الناس يقبلون بضرورة قيام قانون حق المؤلف بتحقيق توازن بين حق العامة في الوصول إلى المعلومات، والحوافز الاقتصادية لمؤلفي مصادر المعلومات وناشريها، غير أنه ليس هناك إجماع بينهم على كيفية هذا التوازن أو طبيعته، ثم أن هناك متشددين من كلا الفريقين؛ حيث حاول أصحاب المصالح بدافع الخوف والطمع أن يستصدروا تشريعاً يخدم مصالحهم فقط، ولسنوات عديدة قدمت عدة قرارات للكونجرس لتغيير قانون حق المؤلف أو تعديله فيما يعالج قضايا المعلومات المباشرة على الإنترنت وبشكل يخدم مصالح مجموعة بعينها، وقد

تضامنت بعض المؤسسات التجارية فيما بينها وسعت إلى امتلاك حقوق متطرفة أو قوية تسيطر من خلالها على المعلومات، وتفرض بموجبها عقوبات على جميع الأنشطة التي ليس لها تصريح رسمي وكأنها أشبه بجريمة وعلى الجانب الآخر ذهبت الجماعات التي تدافع عن مصالح العامة إلى أنه ليس هناك دليل على أن الاستخدام العادل يلحق الضرر بمصالح أي طرف، كما أن القيود التي لا مبرر لها على تدفق المعلومات لها آثار سلبية بالغة. وحتى عام ١٩٩٨م ظل الموقف على حاله متأزماً بين الفريقين، وربما كان في ذلك مصلحة، فقد كان التشريع الساري ملائماً لظهور أولى مراحل النشر الإلكتروني والمكتبات الرقمية، وكانت المشكلة الرئيسة متجسدة في فهم القضايا الكامنة وراء ذلك، صحيح كانت هناك حاجة لإيضاح قانوني في آخر الأمر، إلا أنه كان من الأفضل مراقبة الأنماط التي ظهرت بدلاً من التسرع باستصدار تشريع مبتسر أو مولود قبل أوانه، وربما يكون في التشريع الصادر عام ١٩٩٨م والموضح في اللوحة رقم (٦-٢)، ما يكفي للسماح للمكتبات الرقمية بالازدهار.

اللوحة (٦- ٣)

أحداث في تاريخ حقوق التأليف

يعد مكتب حقوق التأليف خارج الولايات المتحدة سلسلة من اللوحات التي تلخص بعض الأحكام القانونية عن قانون حقوق التأليف، والتي كانت المحاكم الفيدرالية بما في ذلك "المحكمة العليا" قد سنتها. وتُظهر هذه اللوحات كيفية إسهام الشواهد أو السوابق القانونية عبر السنين بصياغة القانون وإيضاحه والسماح بالتطرق إلى مجالات لم تكن تحظى بالاهتمام وقت كتابة الدستور وسن القوانين مثل قضايا التصوير والإذاعة والحاسبات الآلية. وفي الفقرات التالية ملخص لبعض من هذه الأحكام والشواهد:

- **دعـوى هـيوتون ضـد بيترز Wheaton Vs. Peters (١٨٣٤م):** هذه الحالـة التـي تمـثل نقطـة تحـول فـي تـاريخ حقوق التأليف رسـخت مـبدأ مفـاده أن حقـوق التأليف ليست نوعاً من أنواع الحقوق المادية، بل هي استحداث لقانون حقوق التأليف، وهي تخضع للظروف التي فرضتها.

- **دعـوى بيكـر ضـد سـيلدين Baker Vs. Selden (١٨٨٠م):** وقـد رسـخت هذه الحالـة مبدأ أن قانـون حقوق التأليف يحمـي مـا يكتبه أي مؤلـف، ويحمـي طـريقته فـي التعبير عن أفكاره، لكنه لا يحمي الأفكار نفسها.

- **دعوى شركة بورو ـ جيلز ضد شركة ساروني Giles Lithographic Co. Vs. Sarony (١٨٨٤م):** وقد وسـع هذا القرار نطاق قانون حقوق التأليف وجعلهـا تشمل وسـائل أخرى غير النص مثل صورة أوسكار وايلد Oscar Wilde موضع النزاع بين هاتين الشركتين.

- **دعـوى شـركة بليـستين ضـد شـركة دونالدسـون Bleistein Vs Donaldson Lithographic Co. (١٩٠٣م) :** وقـد دارت هـذه القـضية حـول ثلاثـة ملصقات لإعلانات السيرك، وقد قضت المحكمة بأن جميعها تخضع للحماية القانونية سواء أكانت ذات قيم فنية أو كان القصد منها هو المتعة الجمالية.

- **دعوى شركة فريد فيشر ضد شركة ديلنجهام Fred Fisher Inc. Vs. Dillingham (١٩٢٤م) :** وتـناول هذا النـزاع أوجه الشبه بين قطعتين موسـيقيتين، وقضت المحكمة بأن النسخ غير المقصود لقطعة موسيقية

قد يؤدي إلى الإضرار بحقوق التأليف وانتهاكها.

- **دعوى شركة نيكولز ضد شركة يونيفرسال السينمائية Nichols Vs. Universal Pictures Corp. (١٩٣١م) :** وقضت المحكمة في هذا النزاع بأن الأمر ليس فيه انتهاك لحقوق التأليف عندما يقتبس أحد الأفلام أفكاراً مجردة تخص حبكة وشخصيات مسرحية ناجحة تعرض على مسارح برودواي.

- **دعوى شيلدون ضد شركة مترو جولدين السينمائية Sheldon Vs. Metro-Goldwyn Pictures Corp. (١٩٣٦م):** وقد قضت المحكمة في تلك القضية بأنه " ليس هناك منتحل يستطيع أن يبرر الخطأ بإظهار القدر الذي لم ينتحله من إنتاج غيره".

- **دعوى شركة ريكوردي ضد شركة باراماونت السينمائية G. Ricordi Co. Vs. Paramount Pictures Inc. (١٩٥١م)** : وكانت هذه القضية تدور حول كيفية تفسير حقوق التجديد وحقوق تأليف الأعمال المقتبسة من أعمال أخرى مثل رواية "مدام بترفلاي Madame Butterfly" لمؤلفها جون لوثر لونج John Luther Long، والتي اقتبست منها مسرحية لبيلاسكو Belasco، ثم جاء بوشيني Puccini ليقتبس من المسرحية أوبرا جديدة، وقضت المحكمة بأن حماية حقوق التأليف للأعمال المقتبسة من غيرها تنسحب فقط على المادة الجديدة المضافة على العمل القديم.

- **دعوى شركة الإخوة وارنر السينمائية ضد شركة كولومبيا لنظم**

البث الإذاعي Warner Bros. Picture Inc. Vs. Colombia Broadcasting System Inc. (١٩٥٥م): وقد قضت هذه الحالة بأن شخصية سام سبيد Samp Spade في قصة "مالتيس فالكون أو صقر مالطة Maltese Falcon" كانت أداة من أدوات القصة وليست عنصراً من عناصر العمل يخضع للحماية القانونية من قبل قانون حق التأليف.

- **دعوى ميزر ضد شتاين Mezer Vs. Stein (١٩٥٤م):** وقضت المحكمة بأن قانون حقوق التأليف لايحمي الكائنات المفيدة أو النافعة Utilitarian objects، وكانت هذه القضية تدور حول إحدى المصابيح المنحوتة Sculptural Lamp، وإن كان من الممكن تسجيل الخصائص التصويرية والنحتية الخاصة بقطعة نافعة.
- **دعوى كينج ضد شركة السيد مايسترو King Vs. Mr. Maestro Inc. (١٩٦٣م):** ودارت هذه القضية حول خطاب مارتن لوثر بعنوان "أنا أحلم بـ"، ورغم أنه ألقى الخطاب أمام حشد كبير مذاعاً على الهواء مباشرة على الراديو والتليفزيون، فإن المحكمة قضت بأن هذا الخطاب العام لا يشكل عملاً مطبوعاً، ويمكن تسجيل هذا الخطاب في مكتب حق المؤلف باعتباره عملاً غير منشور.
- **دعوى مؤسسة ليتر إد جيد إن بلاك بريس ضد لجنة شيكاجو للمباني العامة Letter Edged in Black Press Inc. Vs. Public Building Commission of Chicago (١٩٧٠م) :** ودارت هذه القضية حول عرض تمثال لبيكاسو على الجمهور، وقد تجاوزها تشريع لاحق.

- **دعوى شركة وليامز للإلكترونيات ضد شركة أرتك إنترناشيونال Williams Electronics Inc. Vs. Artic Internctional Inc. (١٩٨٢م) :** ودارت هذه القضية حول نسخ لعبة من ألعاب الفيديو. وقضت فيها المحكمة بأن مكونات لعبة الفيديو تسري عليها الحماية القانونية. ويمكن اعتبار ذاكرة الحاسب المخصصة للقراءة فقط نسخة [من هذه المكونات].
- **دعوى شركة نوريس للصناعات ضد المؤسسة الدولية للبرق والهاتف Norris Industries Inc. Vs. International Telephone and Telegraph Cor. (١٩٨٣م):** وقد قضت المحكمة بأنه حتى لو رفض مكتب حق التأليف عملاً ما لأنه لا يخضع لقانون حقوق التأليف، فإن صاحبه له الحق في رفع دعواه أمام المحاكم.

حتى وإن اعتُبرت هذه القرارات الجوهرية غير نهائية، فإن العديد منها لم تُدرس أو تختبر من قبل المحكمة العليا وهناك احتمال بإلغائها. ومنذ وقت قريب فقط قامت المحكمة الفيدرالية بوضع قاعدة صريحة بعدم قبول الحكم في دعوى كينج ضد شركة السيد مايسترو King vs. Mr. Maestro.

الخصوصية privacy :

هناك في الولايات المتحدة على الأقل شعور قوي لدى القائمين على المكتبات بأن المستفيدين لهم الحق في احترام خصوصيتهم، وليس من حق أي إنسان معرفة دخول غيره على كتاب معين يتناول مسائل ذات طابع حساس كمرض معين مثلاً، وقد فضلت بعض المكتبات أن تدخل في منازعات قضائية على إفشائها قراءة أحد روادها كتبًا في الشيوعية أمام بعض أجهزة الشرطة،

وتطبق كثير من الولايات قوانين تمنع المكتبات من جمع معلومات تنتهك خصوصية روادها، وهي الأعراف نفسها التي تسري على عالم الإنترنت، ورغم أن المؤسسات لها الحق القانوني في مراقبة نشاط موظفيها، فإن معظم من يعملون في وظائف فنية لديهم الاعتقاد أن بريدهم الإلكتروني وملفات حاسباتهم لابد أن تعامل بخصوصية في معظم الظروف العادية.

لكن كثيراً من تقنية المكتبات الرقمية تستخدم في أغراض تجارية إلكترونية، فشركات الإعلان والدعاية تعمل جاهدة في جمع المعلومات عن عملائها (غالباً دون علمهم بذلك)، ثم تبيع بعض هذه المعلومات لبعضهم، كما أن التقنيات المستخدمة لأغراض معينة مثل تسجيل دخول مستفيد على أحد المواقع، يمكن استخدامها وسيلة لتتبع سلوكه دون علمه.

وكما هو وارد في اللوحة رقم (٦-٤)، ينبغي على المكتبات الرقمية أن تعمل على جمع معلومات عن مدى الإفادة منها، لأن المعلومات الجيدة لازمة لإدارة النظم الآلية وضبطها، وللتكهن بالمشكلات المتوقعة، ووضع خطط التطوير. وبشيء من الحرص يمكن تجميع الإحصاءات عن مدى الاستفادة دون تحديد أشخاص بعينهم، غير أن الحرص لا يتوافر لدى الجميع، فعند معالجة أي مشكلات في شبكة الحاسبات، قد يضطر مدير النظام إلى الاطلاع على أي ملف بالحاسبات أو فحص كل رسالة تمر عبر الشبكة. وقد يحدث قدراً أن يصادف معلومات شديدة الخصوصية أو أعمالاً يجرمها القانون، إذن ما التصرف الصحيح في مثل هذه الأحوال ؟ وما الرأي القانوني في هذه الحالات ؟

اللوحة رقم (٦-٤)

المكتبات الرقمية والخصوصية

لا شك أن كل من يقدم خدمة عليه أن يتزود بالمعلومات الكافية عن كيفية الإفادة من هذه الخدمة، ولإدارة نظام آلي كالمكتبة الرقمية يتطلب الأمر معلومات عن أداء النظام ومدى الاعتماد عليه، ومدى الإفادة من إمكانياته، وعن اتجاهاته وأقصى ما يمكن أن يستفاد منه. وتحتاج المكتبات والناشرون إلى بيانات إحصائية عن كيفية الإفادة من المجموعات ومدى انتشار بعض الخدمات، كما يعتمد من يقومون بتصميم واجهات التعامل Interfaces على معرفة كيفية تعامل المستفيدين مع المكتبة، وعليه يتطلب أمن المكتبات الرقمية حرصاً في تتبع أنماط الإفادة وتحليل درجات الانحرافات عن المعدلات السنوية.

وتجدر الإشارة إلى أنه ليس من السهولة جمع هذه المعلومات، وإن كانت مواقع الإنترنت تقوم بتجميع بيانات عن كيفية الوصول إلى كل ملف. ومن الأساليب الشائعة في ذلك إحصاء مرات الدخول على المواقع، ونظراً لأن كل صورة في شكل رسم هي في العادة عبارة عن ملف مستقل، فإن المستفيد الواحد قد يقرأ صفحة واحدة يمكن أن يدخل عليها عدة مرات. وهذا الإحصاء يفيد في وضع طريقة للتعامل مع النظام الآلي فقط وليس أكثر من ذلك، وقد يرغب مدير إحدى المجلات الإلكترونية المتاحة على الإنترنت في التعرف إلى عدد القراء الذين يقرؤون المجلة عن طريق البيانات المتوافرة عن الإفادة من مقالات بعينها. وهكذا فإن عدد المرات التي يدخل فيها القراء على صفحة المحتويات تعد معلومة مفيدة، وكذلك الحال مع عدد مرات الدخول على الصفحة الأولى من كل مقالة، وهذه مقاييس غير مباشرة لا تفرق بين من يقرأ إحدى القصص عدة مرات وبين عدد من يقرأ تلك القصة مرة واحدة. صحيح

أن هذه الوسائل تخبرنا عن عدد مرات الدخول على النظام الآلي، لكنها تعجز عن تقدير عدد القراء الذين يدخلون على نسخ من المواد من خلال المواقع المطابقة أو المخبأة Mirror or Caches. وإذا كان من الممكن التعرف إلى المستفيدين وتحديدهم، يصبح من السهولة البالغة تجميع معلومات عن سبل إفادتهم وكذلك تجميع إحصاءات دقيقة عن ذلك، لكن تبقى الخصوصية هي المشكلة.

وقد تعاملنا مع هذه المشكلة في "جامعة كارنيجي ميلون" على النحو التالي: في كل مرة تم فيها وصول المستفيد إلى مادة بعينها، يتم إنشاء تسجيلة مؤقتة تحتوي على نسخة مختلطة Scrambled من هوية حاسب المستفيد، ثم يتم الاحتفاظ بالتسجيلات المؤقتة في جزء خاص من نظام الملفات لم ينسخ مطلقاً في شرائط احتياطية. وفي كل أسبوع كانت التسجيلات تحلل للخروج بتقرير ليس الهدف منه تحديد كل مستخدم على حدة، بل تجميع معلومات عن عدد جميع القراء الذين بحثوا كل مجموعة من المواد، وكذلك عدد مرات البحوث التي أجريت، ثم يتم بعد ذلك التخلص من كل التسجيلات المؤقتة هذه. وكذلك من خطوات تتبع جميع هويّات حاسبات المستفيدين.

براءات اختراع البرمجيات :

بالرغم من قصور القوانين وما يكتنفها من عيوب، فإنها قد أحسنت التعامل بشكل عام مع النمو السريع لعالم استخدام الحاسبات الآلية وشبكة الإنترنت، غير أن الاستثناء الأسوأ هو ما يتعلق ببراءات اختراع البرمجيات، فهناك بعض مجالات القانون لا تزال بعيدة كل البعد عن واقع المجال الذي

تطبق فيه، بل قد يحدث في كثير من الحالات أن يقر مكتب براءات الاختراع في الولايات المتحدة بعض البراءات التي تعد تافهة من وجهة نظر المتخصصين في الحاسبات، وحتى وقت قريب لم يكن بمكتب براءات الاختراع علماء حاسبات مؤهلون وقادرون على تقييم الجانب التطبيقي لبعض الاختراعات، ولا يزال المحكمون يمنحون براءات في غاية العمومية، وتتناول أفكاراً ومفاهيم لم تكن خافية على الجميع منذ سنين، أو أنها تصلح للتطبيق في مجالات أبسط مما هو معروف ومتفق عليه في مجالها.

ويكمن أحد أسباب تلك الفوضى في أن مفهوم الاختراع الذي يستند إليه قانون براءات الاختراعات - كما حدث مثلاً مع قفزة أرشميدس من الماء صائحاً: "وجدتها"- لا يتناسب وعلوم الحاسب الآلي، مع أن هناك أعداداً هائلة من الأفكار الجديدة في عالم البرمجيات، ومجتمع علوم الحاسب مجتمع في غاية التجانس، والمنتسبون لهذا المجتمع عادة ما يتلقون تعليمهم في الجامعات نفسها تقريباً، كما أنهم يستخدمون الحاسبات والبرمجيات نفسها تقريباً، وهناك تبادل غير محدود للأفكار عبر العديد من قنوات الاتصال، مما ينتج عنه أن تتعامل مجموعات متشابهة من الباحثين مع المشكلات نفسها، وتتبع من الأساليب الفنية القياسية نفسها بالطرق المتعددة ذاتها.

ولقد كان هذا النمط من التبادل غير المحدود للأفكار من أهم العوامل الكامنة وراء نجاح الإنترنت والانتشار السريع للمكتبات الرقمية، لذلك فإن قانون براءات الاختراع بما فيه من تركيز على السرية والتقاضي والمواجهة لن ينجم عنه سوى الإضرار بهذه الأنماط.

الفصل السابع

إدارة الوصول للمعلومات[1] وأمن المعلومات

يتناول هذا الفصل موضوعين مرتبطين بعضهما ببعض، وهما: طرق ضبط عملية التداول من قبل من لهم حق الوصول إلى مقتنيات المكتبات الرقمية، وأساليب الأمن التي تتوافر في شبكات الحاسبات الآلية.

وتعد الجوانب الاقتصادية من أهم مبررات ضبط عملية الوصول للمعلومات، ومن ثم فإن الناشرين الذين يتوقعون تحقيق عائد مادي مما ينتجون من مواد، يمنحون حق الوصول للمعلومات فقط للمستفيدين الذين دفعوا الرسوم المطلوبة. وقد يرى البعض أن لا حاجة لفرض قيود على إدارة الوصول للمعلومات إلا عندما يكون العائد المادي هو الأساس. غير أن هناك أسباباً أخرى تدعو إلى ضبط عملية تداول مقتنيات المكتبات الرقمية، فقد يقيد استخدام بعض المواد بشروط معينة يفرضها الذين أهدوا هذه المواد (كأن يقيد الاستخدام بحياة هؤلاء أو وفاتهم)، كما قد لا ترغب بعض الجهات في تداول مجموعاتها لما تتضمنه من معلومات ذات طابع سري، كالأسرار التجارية وسجلات الشرطة أو وثائقها، والمعلومات الحكومية المصنفة

(١) يستخدم مصطلح "إدارة الوصول للمعلومات" في هذا الكتاب ليصف عملية التحكم في الوصول إلى مقتنيات المكتبات الرقمية وتداولها، وإن كانت هناك مصطلحات أخرى تشير إلى المعنى نفسه، مثل "الضوابط والشروط"، ومن الملاحظ أن ثمة مصطلحًا غريبًا يشيع استخدامه في مجال النشر حيث التركيز على العائد المادي، وهو مصطلح "إدارة الحقوق". وبالرغم من عدم وجود تفاوت واضح بين هذه المصطلحات، فإن لكل واحد منها جانبًا من التركيز لا يتوافر لغيره.

classified government information. كما أنه ليس من السهولة وضع حدود فاصلة بين عوالم الفن والإسفاف وانتهاك خصوصية الآخرين، وحتى عندما يكون تداول المجموعات أو الوصول إليها مفتوحاً أو متاحاً، فإن أساليب ضبط عمليات الإضافة والتغيير والحذف لمحتويات هذه المجموعات وما وراء البيانات الخاصة بها، يعد مطلباً له ما يبرره. وتحرص الإدارة الجيدة للمكتبة الرقمية على الاحتفاظ بسجل يضم جميع التغيرات حتى يمكن إعادة اختزان تلك المجموعات في حالة حدوث أية أخطاء أو تلف يصيب ملفات الحاسب الآلي.

إن إدارة الوصول للمعلومات يحيطها شيء من عدم وضوح الرؤية، حيث يفترض من لديهم خلفية في الحاسب الآلي أنه من الممكن في بعض الأحيان إن يتم عنونة كل كائن رقمي بمجموعة ما وراء البيانات التي تضم قائمة حقوق الاستخدام وتصاريحه، والعوامل الأخرى المتصلة بإدارة الوصول للمعلومات. هذا في الوقت الذي يعلم فيه من لهم خلفية في مجال المكتبات، وخاصة المسؤولين عن المجموعات التاريخية والأرشيفية، أن تجميع مثل هذه المعلومات عادة ما يكون مضيعة للوقت، بل أحياناً ما يكون ضرباً من المستحيل. إن مشروعات كمشروع الذاكرة الأمريكية الذي تتبناه مكتبة الكونجرس، تسعى إلى تحويل الملايين من المواد من المجموعات التاريخية، ويبدو من الطبيعي في حالة تلك المواد القديمة افتراض انتهاء فترة حماية حقوق التأليف الخاصة بها، وبالتالي إزالة كل القيود على تداول هذه المواد، غير أن ذلك أبعد ما يكون عن الحقيقة. فانقضاء فترة حماية حقوق التأليف للمواد المنشورة مرتبط بوفاة صاحبها، وهو أمر يصعب تحديده أو التكهن

به، كما أن المكتبات غالباً لا تعرف ما إذا كان قد تم نشر مادة بعينها أم لا.

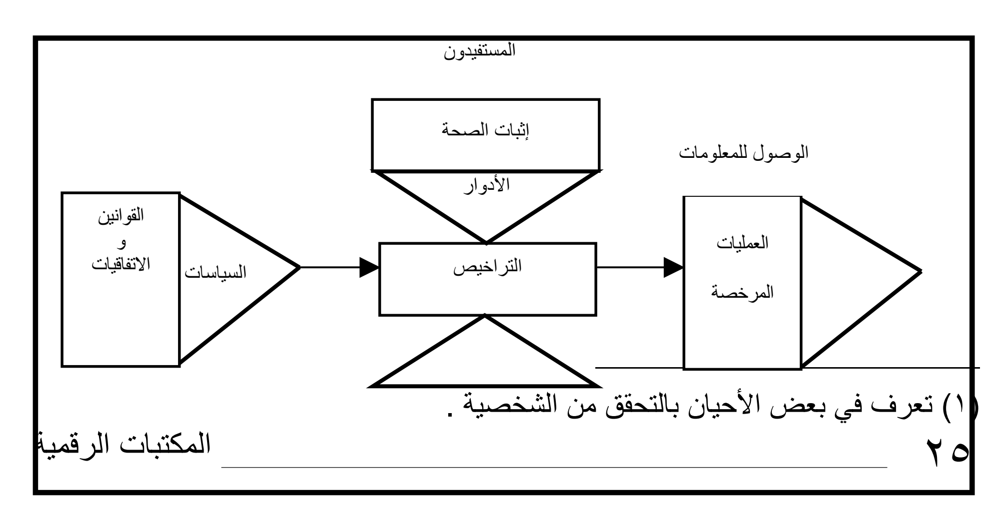
وكما أوضحنا في الفصل السادس، فإن كثيراً من القوانين التي تحكم عمل المكتبات الرقمية، مثل قوانين حقوق التأليف، ليس لها حدود واضحة، ومن ثم فإن السياسات الخاصة بإدارة الوصول للمعلومات، والتي تستند إلى تلك القوانين يحيطها الغموض نفسه، وإذا قدر أن تتضح هذه المعالم من خلال ما يستجد من قوانين أو معاهدات أو إجراءات قانونية، فإن تلك السياسات ستتغير تبعاً لذلك.

عناصر إدارة الوصول للمعلومات :

يوضح الشكل رقم (٧-١) نظامًا يفيد في عملية التفكير بإدارة الوصول للمعلومات؛ حيث يظهر على اليمين من الشكل مديرو المعلومات الذين يضطلعون بمهام وضع السياسات الخاصة بالوصول للمعلومات التي تربط بدورها المستفيدين (في قمة الشكل) بالمواد الرقمية (في أسفل الشكل)، هذا في الوقت الذي يتمركز فيه الترخيص بالاستخدام (في وسط الشكل) محدداً أساليب الوصول للمعلومات (على اليسار من الشكل). ويحتاج كل جانب من هذه الجوانب إلى إيضاح؛ فلا بد أن يضع مديرو المعلومات في اعتبارهم عند وضعهم للسياسات والقوانين المتصلة بذلك، الاتفاقيات الموقعة مع الآخرين (مثل التصاريح الممنوحة من أصحاب حقوق التأليف)، كما يجب التثبت من هويات المستفيدين أو شخصياتهم، وأن يحدد دورهم في الوصول لهذه المواد، كذلك يجب أن تعرف المواد الرقمية في المجموعات أو تحدد، فضلاً عن التثبت من هوية تلك المواد، وعلى الطرف الآخر فإن الوصول لهذه المواد لا يتم إلا عبر العمليات المرخص بها.

وعندما يطلب المستفيد إتاحة فرصة الوصول لمادة رقمية، فإن طلبه يمر عبر إجراءات عملية إدارة الوصول للمعلومات، وبعد التثبت من شخصية المستفيد من خلال عمليتي "التثبت من الهوية [١] Authentication والترخيص بالاستخدام Authorization"، تقوم هذه الإجراءات بمنح هذا المستفيد حق تنفيذ عمليات محددة أو منعه من ذلك.

إن مسؤولية التعامل تقع عادة على الطرف الذي يتولى إدارة المواد الرقمية، وقد يكون هذا المدير مكتبة ما، أو ناشراً، أو مدير أحد المواقع على مدير الويب webmaster، أو منتج المعلومات نفسها، ويمكن أن يفوض جزءًا من هذه المسؤولية لطرف آخر. وإذا كان هذا الطرف إحدى المكتبات التي تتحكم في المواد وتتيحها للمستفيدين، فإنها تضع السياسات وتعمل على تطبيقها، وعادة ما يحكمها في ذلك القيود الخارجية، كالضوابط القانونية، أو التراخيص الممنوحة من قبل الناشرين، أو الاتفاقيات مع المتبرعين. وإذا قام أحد الناشرين بإتاحة المواد، ورخص بتداولها، فإن هذا الناشر يصبح هو المسؤول عن إدارتها (المدير)، ويمكنه في هذه الحالة أن يفوض آخرين في صلاحيات الأنشطة الأساسية كالترخيص للمستفيدين.



(١) تعرف في بعض الأحيان بالتحقق من الشخصية .

التعريف والتثبت في الهوية
المادة الرقمية

الشكل رقم (٧ - ١) إدارة الوصول للمعلومات

المستفيدون :

التحقق من الهوية Authentication :

ثمة أساليب كثيرة ومتنوعة تستخدم في التثبت من هويات المستفيدين، منها ما يتميز بالبساطة - وإن كان يعيبه سهولة الاختراق، ومنها ما يتمتع بقدر كبير من الأمان، لكن يعيبه التعقيد. ويمكن تقسيم هذه الأساليب إلى أربع فئات رئيسة على النحو التالي :

ـ ما الذي يعرفه المستفيد؟

من الأساليب القياسية وشائعة الاستخدام للتثبت من هويات المستفيدين، أن يمنح كل مستفيد اسمًا مميزًا للدخول وكلمة سر. وبالرغم من سهولة هذا الأسلوب فإنه عرضة لإساءة الاستخدام، كما أن من السهولة سرقة كلمات السر؛ لأن المستفيدين غالباً ما يختارون كلمات سهلة التذكر، ومن ثم تخمينها بسهولة.

ـ ما الذي يمتلكه المستفيد؟

هناك نوعان من بطاقات التثبت من هوية المستفيد، هما: البطاقات المشفرة مغناطيسياً magnetic encoded cards كالتي تستخدم في أجهزة الصراف الآلي، "والبطاقات الذكية الرقمية digital smartcards" التي تنفذ برنامج التثبت. وبشكل عام تعد هذه البطاقات الذكية من أفضل نظم التثبت،

حيث تتمتع بدرجة عالية من الأمان فضلاً عن سهولة الاستخدام.

ـ أين يتواجد المستفيد؟

من نماذج التثبت شائعة الاستخدام : عنوان الحاسب على الشبكة، حيث يمكن التثبت من أي شخص لديه إمكانية الوصول لحاسب معين له عنوان معتمد مبني على بروتوكول الإنترنت، وعلى ذلك فإن البيانات المحملة على كثير من الحاسبات الشخصية لا تتوافر لها الحماية عادة إلا من قبل صاحب الجهاز، أي أن كل من له الحق في استخدام الحاسب يمكنه قراءة تلك البيانات.

ـ ما السمات الشخصية للمستفيد؟

يستخدم أسلوب التثبت من هوية المستفيد عن طريق سماته الشخصية، كنبرات الصوت على سبيل المثال، في القليل من التطبيقات السرية، غير أن استخدام هذا الأسلوب في مجال المكتبات الرقمية لا يزال ضعيفاً.

الأدوار Roles :

نادراً ما تحدد سياسات إدارة الوصول للمعلومات المستفيدين عن طريق أسمائهم، لأنها مرتبطة عادة بفئات المستفيدين أو بدور كل مستفيد على حدة. وقد يكون للمستفيد عدة أدوار؛ حيث يمكن للمستفيد نفسه أن يستخدم المكتبة الرقمية في أوقات مختلفة لأغراض التدريس أو للقراءة الحرة، أو لأغراض العمل الجزئي غير المتفرغ. وتبعاً لذلك فإن المكتبة الرقمية قد يكون لها سياسات مختلفة تجاه هذا المستفيد نفسه، بحيث تتفاوت بتفاوت كل دور من تلك الأدوار. ومن أهم الأدوار في هذا السياق ما يلي :

- **عضوية إحدى المجموعات:** فقد يكون المستفيد أحد أعضاء المعهد (الأوروبي) للفيزياء على سبيل المثال، أو قد يكون طالباً في الأكاديمية البحرية الأمريكية.

- **موقع المستفيد:** فقد يستخدم المستفيد حاسباً في مكتبة جامعة كارنيجي في بتسبرج، أو في مكان ما من نيوزيلندا.

- **مدى اشتراك المستفيد:** فقد يكون المستفيد مشتركاً اشتراكاً ساري المفعول في مجلة جمعية حرفيي استخدام الحاسبات، أو أحد منسوبي إحدى الجامعات التي تمتلك الترخيص لاستخدام جميع مجموعات جستور JOSTR.

- **برنامج التكشيف الآلي:** فقد يكون المستفيد أحد مستخدمي برنامج متصفح الويب أو أي برنامج آلي آخر.

- **أسلوب تسديد الرسوم:** فقد يكون للمستفيد حساب ائتمان مع خدمة ليكزيس Lexis[١]، وأنه ممن يدفعون عشرة دولارات مقابل الوصول للمواد.

إن معظم مستخدمي- أو بالأحرى مستفيدي- المكتبات الرقمية هم عناصر بشرية يستخدمون حاسبات شخصية، إلا أن المستفيد قد يكون حاسباً آلياً يعمل بدون شخص يشغله، أو أحد برامج تكشيف صفحات الويب، أو

(١) إحدى الشركات الرائدة في تقديم المعلومات القانونية النصية الكاملة على الخط المباشر (المترجمان).

أحد برامج المطابقة Mirroring [١] التي تقوم باستنساخ المجموعة بأسرها أو تكرارها، وهناك من المواقع ما يحظر ضمنياً إتاحة الوصول عن طريق برامج آلية أو عن طريق منحها صلاحيات عالية higher privileges.

المادة الرقمية :

تعريف المواد والتثبت من هوياتها:

يجب تعريف المواد الرقمية بوضوح من أجل تحقيق متطلبات إدارة الوصول إليها وتداولها، ويتم ذلك من خلال إعطاء كل وحدة من وحدات تلك المواد اسمًا أو محددًا معينًا. وهذا موضوع رئيسي بالنسبة لكل من المكتبات الرقمية والنشر الإلكتروني.

والتثبت من هوية المواد الرقمية يضمن لكل من المستفيد ومدير المجموعات عدم حدوث أي تغيير في مضمون تلك المواد. وهي مسألة بالغة الأهمية في بعض المواقف؛ فقد سبق لي أن عملت مع بعض الزملاء لدى وزارة التجارة الأمريكية في تجميع عدد من الوثائق ذات العلاقة بالشؤون الخارجية كالمعاهدات والاتفاقيات التجارية، وبدت لنا هنا مدى أهمية دقة الصياغة في تلك الوثائق؛ فلو أن وثيقة ما- على سبيل المثال- ذكر أنها هي التي تتضمن النص الدقيق لاتفاقية التجارة الحرة لأمريكا الشمالية، لوجب على كل من يتعامل مع هذه الوثيقة أن يثق في صحة صياغتها ودقة نصها، وكم من مرة أثارت نصوص غير دقيقة في صياغتها- سواء أكان ذلك عن قصد أو عن غير قصد- أزمات دولية.

(١) نظام آلي يحتوي على نسخة مكررة من المعلومات المختزنة في نظام آخر (المترجمان).

ويحدث في معظم المكتبات الرقمية ألا يتم التثبت من مدى صحة موادها بشكل دقيق، وحيثما يكون مستوى الثقة عالياً وتكلفة الأخطاء قليلة، فليس هناك حاجة أو مبرر لعملية رسمية للتثبت من الهوية، وإذا كانت التغييرات المتعمدة قد تبدو نادرة، كما أن الأخطاء عادة ما يكون من السهولة كشفها. فإن وجود تلك الأخطاء يعد أمراً خطيراً، وخاصة في بعض المجالات الحيوية كالطب على سبيل المثال، وبناء على ذلك فإن على المكتبات الرقمية في مثل هذه الحالات، أن تعيد النظر بشكل جاد في استخدام الطرق الرسمية للتثبت من هوية موادها.

ولضمان دقة كائن معين، فإنه يمكن ربطه بتوقيع رقمي خاص بها (سنتناول في نهاية هذا الفصل الأساليب الفنية للقيام بذلك)، بحيث يكفل هذا التوقيع عدم تغيير أي ملف، أو أية أجزاء من مكوناته بعد وضعه. وتبين اللوحة رقم (٧-١) كيفية استخدام التوقيعات الرقمية في مكتب حماية حقوق التأليف الأمريكي.

وثمة مجموعة من وسائل التثبت من هوية المواد يشار إليها بالمصطلح "الدمغ بالعلامات المائية Watermarking"، وهي وسائل دفاعية يستخدمها الناشرون لتتبع عمليات النسخ غير القانونية وردعها. وتكمن الفكرة الأساسية لهذه الأساليب في دمغ رمز (أو علامة) ما في المادة بطريقة دقيقة ماكرة لا يدركها المستفيد، ولكن يمكن استعادتها لإثبات الملكية تماماً مثلما تفعل شركات الإعلان عندما تضيف شعاراً لإعلان تليفزيوني ليكشف مصدر الصورة إذا ما تم تقليدها. وقد يكون من المستحيل على المستفيد كشف

العلامات المائية الرقمية، لكن من المؤكد أن إزالتها تكاد تكون مستحيلة، وإذا ما تم ذلك فيصبح ذلك أمراً مكشوفاً للعيان.

اللوحة رقم (٧-١)

التسجيل الإلكتروني والإيداع التزاماً بقانون حقوق النشر

يعد المكتب الأمريكي لحقوق النشر إحدى الإدارات المستقلة التابعة لمكتبة الكونجرس، ولمكتبة الكونجرس - حسب القانون الفيدرالي- حق اقتناء نسختين من كل عمل ينشر في الولايات المتحدة الأمريكية. وفي الوقت الذي نجد فيه أن إيداع الأعمال المنشورة أمر إلزاميّ، نلاحظ أن تسجيل الأعمال الخاضعة للحماية القانونية لم تكن مطلباً في قانون حق المؤلف الصادر عام ١٩٧٨م، رغم أن هناك تشجيعاً لعملية التسجيل لما لها من فوائد كبيرة.

وطريقة التسجيل لأغراض حماية حقوق المؤلف بسيطة وواضحة، إذ يقوم مالك هذه الحقوق بإرسال نسختين من عمله إلى مكتب حق المؤلف مع نموذج التسجيل والرسوم اللازمة، بعد ذلك يتم دراسة الطلب والعمل المقدم، وإذا أقر ذلك صدرت للعمل شهادة تسجيل، بعدها يرسل العمل لمكتبة الكونجرس لكي تقرر ما إذا كان يستحق الإضافة لمجموعاتها، أو وضعه ضمن قائمة التبادل مع المكتبات الأخرى.

وفي عام ١٩٩٣م شرع كل من مكتب حقوق التأليف ومؤسسة مبادرة البحوث الوطنية، في استحداث نظام لتسجيل الأعمال الإلكترونية وإيداعها، عرف باسم كوردس Copyright Office Electronic Registration, Recording and Deposit System (CORDS). ويعكس هذا النظام الإجراءات التقليدية المتبعة، فالطلب نموذج متاح على الويب، بالإضافة إلى نسخة رقمية من

العمل تتاح بشكل آمن على الإنترنت، أما الرسوم فتدفع بشكل منفصل. ويتم استخدام التوقيعات الرقمية في التعرف إلى الطلبات المقدمة لتسجيل حق المؤلف؛ حيث يقدم صاحب الطلب مستخدماً مفتاحاً خاصاً به طلباً موقعاً، إضافة إلى العمل نفسه، والتوقيع الرقمي، والمفتاح العام، والشهادات ذات الصلة بالموضوع. ويعمل التوقيع الرقمي على إبلاغ مكتب حقوق التأليف بوصول الطلب بشكل صحيح وتأكيد صحة بيانات هوية صاحب الطلب. وبناء على ذلك وفي حالة حدوث أي نزاع في المستقبل حول العمل، يرجع إلى التوقيع الرقمي للتثبت من صحة الادعاء ومن العمل المسجل.

سمات المادة الرقمية Attributes of Digital Library :

غالباً ما تتعامل سياسات إدارة الوصول للمعلومات مع المواد المختلفة بطرق وأساليب متنوعة حسب سمات تلك المواد وخصائصها، ويمكن تشفير تلك الخصائص باعتبارها "ما وراء بيانات إدارية administrative metadata" يتم اختزانها مع الكائن نفسه، أو أن هذه الخصائص يمكن أخذها من مصدر آخر. كذلك يمكن لبعض السمات أن تحسب، كما هو الحال بالنسبة لحجم الكائن حيث يمكن قياسه. وفيما يلي بعض الأمثلة على ذلك :

- **التقسيم إلى مجموعات فرعية:** غالباً ما تُقَسَمُ المجموعات إلى مواد متاحة بدون قيود، ومواد مقيدة الاستخدام؛ حيث يفصل الناشرون نصوص المقالات الكاملة عن غيرها من الكشافات والمستخلصات والمواد الإعلانية الأخرى، كما تتيح بعض المواقع على الويب حرية الدخول إليها للجميع، في حين تقيد بعض المواقع الأخرى استخدامها بفئة محددة من المستفيدين كأعضاء المؤسسة التابع لها هذا الموقع.

- **الترخيص والالتزامات الخارجية الأخرى:** قد تقتني المكتبة الرقمية مواد مرخصاً بها من قبل ناشريها، أو مواد تخضع لضوابط وشروط تحكم تداولها، كتلك المواد التي تودع في مكتبة الكونجرس تنفيذاً لقانون حق المؤلف.

- **الخصائص المادية والمؤقتة:** ربما يكون للمكتبات الرقمية سياسات تستند إلى عامل الفترة الزمنية، وهو تاريخ النشر، أو على الخصائص المادية كالحجم، حيث تتيح بعض الدوريات حرية التداول المجاني لبعض مقالاتها وذلك فور صدورها، في حين تطلب ترخيصاً لتداول تلك المقالات بعد مضي فترة زمنية على نشرها.

- **أنواع الوسائط:** قد تعتمد المكتبة الرقمية على سياسات تستند إلى شكل المادة، أو على نوع الوسيط حيث يمكن - على سبيل المثال - أن تعامل الوسائط السمعية المرقمنة والمواد النصية وبرامج الحاسب والصور بأساليب مختلفة بناءً على اختلاف طبيعة كل وسيط من هذه الوسائط.

ومن الضروري تناول خصائص المواد الرقمية بطرق متفاوتة granularity، لأنه لو توافرت لجميع المواد في مجموعة معينة الخصائص نفسها لكان من المناسب وصفها وصفاً واحداً، وعلى النقيض من ذلك تماماً، حيث نجد في بعض الأوقات أن أجزاءً معينة من المواد تنفرد بسمات خاصة بها. كما أن الحقوق المتعلقة بالصور تختلف في الغالب عن تلك الحقوق المتعلقة بالنص الذي تلحق به تلك الصور، ومن ثم فلا بد من التمييز بين

هذين الشكلين، ويمكن لشخص ما أن يهدي إحدى المكتبات مجموعة من الرسائل ويسمح بإتاحتها للتداول باستثناء مواد محددة منها يقيد تداولها بشروط معينة. وبناء على ذلك، فإن على المكتبات الرقمية أن تسعى إلى ربط الخصائص بالمجموعات الكلية، أو بالمجموعات الفرعية، أو بمواد معينة في المجموعة، بل بعناصر محددة في تلك الكائنات.

العمليات :

غالباً ما تحدد أو تقيد سياسات إدارة الوصول العمليات وغيرها من الممارسات المختلفة الأخرى التي يسمح للمستفيد القيام بها حيال مواد المكتبة، وهناك نوعان من العمليات هما:

- **العمليات الحاسوبية:** هناك بعض العمليات التي توصف بمصطلحات حاسوبية، مثل: كتابة البيانات على الحاسب الآلي، أو تنفيذ البرنامج، أو تراسل البيانات عبر الشبكة، أو عرض البيانات على شاشة الحاسب، أو طباعة البيانات أو نسخها من جهاز إلى جهاز آخر.
- **حدود الاستخدام:** حيث يسمح للمستفيد باقتباس مواد معينة من إحدى قواعد البيانات، ولكن لا يسمح له بنسخ قاعدة بيانات بأكملها.

ويمكن ضبط العمليات المشار إليها سابقاً بالوسائل الفنية، لكن هناك كثيراً من السياسات التي قد يقررها مدير المعلومات ويستحيل تنفيذها من الناحية الفنية، ومن هذه السياسات ما يلي:

* **العمل أو الغرض:** حيث يمكن أن يشير تفويض المستفيد إلى سبب

القيام بعملية ما، وتشمل الأمثلة على ذلك الاستخدامات التجارية والتعليمية والحكومية.

* **العمليات الفكرية:** فقد تحدد العمليات الاستخدام الفكري لمادة من المواد، وهنا تظهر أهمية القواعد التي تضبط عملية اشتقاق عمل جديد يقوم على محتوى عمل آخر، أي أن المعايير لابد أن تضع في اعتبارها هدف الاستخدام ومداه.

الاستخدام اللاحق [غير المباشر] Subsequent Use :

ينبغي على نظم إدارة الوصول أن تراعي كلاً من العمليات المباشرة direct operations، والاستخدام اللاحق للمواد. ويقصد بالعمليات المباشرة تلك العمليات التي يتولاها الجهاز الخازن repository، أو أي نظام آلي آخر يعمل كوكيل لمدير المجموعة. أما الاستخدام اللاحق فيشمل جميع العمليات التي يمكن أن تحدث بمجرد خروج المادة عن سيطرة المكتبة الرقمية، ويشمل ذلك جميع سبل النسخ (بدءاً من عمليات استنساخ ملفات الحاسب إلى عمليات تصوير المستندات). ومن منظور فكري يمكن أن يتراوح الاستخدام اللاحق بين عمليات اقتباس أجزاء قصيرة، وإيجاد أعمال قائمة على أعمال سابقة، وانتهاءً بالانتحال الصريح للمحتوى الفكري.

وعندما يتم إرسال مادة معينة، أو جزء من إحدى المواد إلى حاسب شخصي، يصعب من الناحية الفنية منع المستخدم من نسخ ما أرسل إليه أو اختزانه أو توزيعه إلى الآخرين، وهو أمر يشبه عملية تصوير المستندات. وإذا كانت المعلومات متاحة للبيع، زادت احتمالات الاستخدام اللاحق لها. والواقع

أن الناشرين تزعجهم فكرة قيام القراء بتوزيع نسخ من المواد دون موافقة منهم على ذلك، وإذا ما تتبعنا الفكرة في أسوأ احتمالاتها وجدنا أنه لو قام أحد الناشرين ببيع نسخة واحدة من إحدى المواد، ثم حدث أن وزعت تلك المادة عبر الإنترنت توزيعاً موسعاً، فقد ينتهي به الحال إلى عدم بيع نسخ أخرى من العمل غير تلك النسخة الوحيدة الأولى؛ وتحسباً لهذه المخاوف تسمح المكتبات الرقمية غالباً لقرائها بتداول تسجيلات فردية أو استخدامها، دون تقديم أية وسائل تكفل لهم نسخ مجموعات كاملة، ومع أن هذا لا يحول دون ضياع قدر ضئيل من العائد المادي من صاحب المادة، فإنه يقف عائقاً أمام كل من يحاول تقويض المصالح الاقتصادية للناشر من خلال نسخ العمل برمته.

السياسات :

يشير التعريف الرسمي لكلمة "السياسة policy" إلى أنها القاعدة التي سنّها مدير المعلومات لتحدد من له حق القيام بعمل أو فعل معين تجاه مادة معينة، ومن السياسات الشائعة في المكتبات الرقمية ما يلي :

- أن يكون لأحد المطبوعات سياسة تحدد إتاحتها؛ بمعنى أن يتاح لأي شخص إمكانية قراءة المادة فقط، في حين يخول أعضاء تحرير هذا العمل الحق في إجراء أي تغييرات عليه.

- أن يكون لناشري الدوريات على الخط المباشر سياسة تسمح للمشتركين دون غيرهم بالوصول إلى هذه الدوريات أو قراءتها، في حين لا تسمح لغيرهم إلا بقراءة صفحة المحتويات، أو المستخلصات، بمعنى أن حق التعامل مع الدورية بمحتوياتها كاملة لا يخول إلا لمن يدفعون رسوماً نظير

استخدامهم تلك الدوريات.

- تقوم بعض الجهات الحكومية بتصنيف المواد (مثل المواد السرية للغاية)، ويكون لها سياسات مشددة حول من له حق الوصول لهذه المواد وتداولها، والظروف التي يسمح له فيها بذلك، وطبيعة ذلك التداول.

وتجدر الإشارة إلى أن تلك السياسات نادراً ما تكون بالبساطة التي عرضنا لها في الأمثلة السابقة، فمجلة المكتبات الرقمية D.L. magazine على سبيل المثال، لها سياسة الإتاحة الحرة للجميع، غير أن مؤلفي مقالات هذه المجلة يحتفظون بحقوق تأليف مقالاتهم هذه. أي أن سياسة الوصول هذه تسعى إلى تشجيع الجميع على قراءة المقالات وطباعتها للاستخدامات الشخصية، أما الاستخدامات اللاحقة (كتأليف عمل معتمد على العمل الأصلي، أو بيع نسخ بقصد الربح)، فيتطلب إذناً بذلك من صاحب حق التأليف.

ونظراً لأن سياسات إدارة الوصول قد تكون معقدة، فقد تطلب الأمر إيجاد بعض الوسائل الرسمية لصياغتها بهدف تبادل المعلومات بين أنظمة الحاسب الآلي، وربما يعد أكثر الأعمال شمولاً في هذا الصدد ما قام به "مارك ستيفكMark Stefik" من شركة زيروكس، حيث طرح مارك ما يعرف بـ "لغة حقوق الملكية الرقمية The Digital Property Rights Language"، وهذا العمل يعتبر لغة تعبر عن حقوق استخدام الأعمال الرقمية، وضوابط ذلك، ورسومه. كما تسعى هذه اللغة إلى تحديد خصائص المواد الرقمية، وسياسات تداولها، بما في ذلك الاستخدام اللاحق لها.

وبموجب هذه اللغة يكون بمقدور مدير المجموعة وضع ضوابط وشروطًا لعمليات نسخ هذه المواد وإرسالها ونقلها وطباعتها وغيرها من العمليات الأخرى المشابهة. وتنص اللغة المقترحة على إيجاد رسوم تُحدد بعد ذلك نظير أي عملية، كما أنها تضع تصوراً للروابط مع آليات الدفع الإلكتروني. أما بالنسبة لنظام الرمز المستخدم في هذه اللغة فهو يعتمد على لغة معالجة القوائم Lisp (١)، وهي لغة تستخدم لمعالجة اللغة الطبيعية، ويرى البعض أن من الأفضل لنظام الرموز في المكتبات الرقمية أن يستخدم لغة التهيئة الموسعة XML (٢). غير أن المحك الحقيقي ليس في نظام الرمز بقدر ما هو في مدى فاعلية اللغة ونجاحها في إثبات وجودها واستخدامها على نطاق واسع.

تنفيذ سياسات إدارة الوصول :

ليست عملية إدارة الوصول مجرد مسألة وضع سياسات ملائمة، بل العبرة ـ كما يريد مديرو المعلومات-في تطبيق تلك السياسات، وهو أمر يتطلب شيئاً من الإلزام.

---

(١) (list procssing) Lisp لغة معالجة القوائم، إحدى لغات البرمجة صممها جون مكارثر في عامي ١٩٥٩- ١٩٦٠م، واستعملت أساساً لتناول قوائم البيانات بالتعديل والترتيب، وقد استعملت هذه اللغة على نطاق واسع في دوائر الأبحاث والدوائر الأكاديمية، حيث تعتبر اللغة المعيارية لأبحاث الذكاء الاصطناعي (المترجمان).

(٢) إصدارة مبسطة من اللغة المعيارية الموحدة لتهيئة النصوص SGML، ولغة XML هذه خاصة بالوثائق المعروفة على الويب تساعد المستخدمين على إضافة مهام للوثيقة غير موجودة في لغة HTML (المترجمان).

إن بعض السياسات يمكن تطبيقها بشكل إلزامي من الناحية الفنية، لكن ذلك ليس ممكنًا مع جميع السياسات؛ فهناك مثلاً وسائل فنية واضحة لتطبيق سياسة معينة مع من يصرح له بتغيير مادة ما في مجموعة معينة، أو بالبحث في مستودع الوثائق. لكن ليس هناك وسائل فنية لمكافحة الانتحال، أو انتهاك الخصوصية أو ضمان استخدام المواد للأغراض التعليمية دون غيرها من المجالات. فمثل هذه السياسات- وبرغم ما تتمتع به من بعد منطقي- تبدو غاية في الصعوبة عند تنفيذها بالوسائل الفنية، ولا سبيل إلى تنفيذها إلا بالوسائل القانونية والاجتماعية.

وهناك علاقة عكسية بين دقة التنفيذ ومدى رضا المستفيدين؛ فالوسائل الفنية لتنفيذ السياسات قد تسبب لهم إزعاجاً؛ فقليلون هم أولئك الذين يعترضون على كتابة "كلمة مرور password" عند بداية الدخول إلى الشبكة، لكن الجميع لا يرغبون في إعادة كتابة كلمة المرور أو غيرها من إجراءات أخرى للتحقق من الهوية عدة مرات. وقد يقرر مديرو المعلومات في بعض الأحيان تخفيف قيود عملية تطبيق السياسات من أجل إرضاء المستفيدين، على اعتبار أن إرضاءهم يسهم في نمو السوق حتى ولو ضاع جزء من العائد بسبب المستفيدين غير المصرح لهم بالدخول. وكما يحرص الناشرون الذين يتحمسون كثيراً لتطبيق السياسات على سعادة المستفيدين، وغالباً ما تكون عائداتهم أكبر من عائدات غيرهم من الناشرين. وكما يبدو من اللوحة رقم (٧-٢)، فإن هذه تعد الاستراتيجية المتبعة الآن في الغالبية العظمى من برمجيات الحاسبات الشخصية. وهو التوجه نفسه الذي يعمد إليه بعض ناشري الدوريات الإلكترونية، ومثال ذلك دار نشر هاي وير High wire Press.

وإذا ما أخذ التراخي مأخذه من الوسائل الفنية، فيمكن للضغوط الاجتماعية والقانونية أن تؤدي دورها، ومن بين الأهداف الاجتماعية تدريب المستفيدين على السياسات المطبقة على المجموعات وإقناعهم باتباعها، أو حتى ملاحظتهم بذلك، وتحتاج هذه الأهداف إلى سياسات يسهل فهمها وبالتالي إتباعها. كما ينبغي إحاطة المستفيدين بالسياسات وإعلامهم بطبيعة التصرفات السليمة. ومن الوسائل المفيدة أن تظهر على الشاشة "عبارات إرشادية" بمجرد الوصول إلى المادة، وهذه العبارة بمثابة نصوص تحدد بعض السياسات، مثل: "من أجل الحفاظ على حقوق التأليف" يحظر استخدام هذه المادة لأغراض تجارية. وهناك من الوسائل غير الفنية في مجال تطبيق السياسات ما هو أكثر حزماً، مثل إقدام الناشر على إلغاء الترخيص الذي يمنحه لأعضاء إحدى المؤسسات إذا ما دأبوا على انتهاك اتفاقية الترخيص، أو أساءوا استخدام السياسات التي ينبغي عليهم احترامها والالتزام بها، أما أشد الإجراءات قسوة فهو اللجوء للقضاء، وعندها يمكن القول: إن قضية واحدة يحسن أصحابها الدعاية لها، قد تقنع الكثيرين بالتحلي بالمسؤولية إزاء التعامل مع المواد.

اللوحة رقم (٧-٢)

سياسات إدارة الوصول للبرمجيات

تقدم التجارب المبكرة في برمجيات الحاسبات الشخصية نموذجاً لما يحدث عندما لا تلقى محاولات تطبيق السياسات قبولاً من المستفيدين.

ومن المعتاد أن البرامج يرخص بها لحاسب واحد، كما تغطي رسوم الترخيص استخدام البرمجية في حاسب واحد، لكنه من السهل نسخ مثل هذه

البرمجيات، وعادة ما يخسر منتجو البرامج العائد من هذه البرمجيات المنسوخة، وخاصة إذا تم توزيعها على نطاق واسع.

وقد حاول منتجو البرمجيات في بدايات استخدام الحاسبات الشخصية أن يضبطوا عملية الاستنساخ غير المرخص للبرمجيات بأساليب فنية، منها أن تحمل البرمجيات على أقراص لا يمكن نسخها بسهولة، وهو ما يعرف باسم "الحماية ضد النسخ copy protection"، وفي كل مرة يريد المستفيد تشغيل البرنامج عليه أن يضع الأسطوانة الأصلية، وقد كان لهذا أثره الواضح على سوق النشر، وهو ما لم يكن يطمح إليه منتجو البرمجيات، ولم يكن مناسباً للمستفيدين المرخص لهم باستخدام البرنامج الذي اشتروه، كما كانت عملية تنصيب القرص الصلب وحفظ البرنامج تتسم بالصعوبة، وقد اعترض المستفيدون على ذلك، وكانت النتيجة أن موردي البرامج الذين كانوا من دعاة حمايتها، قد خسروا مبيعات كثيرة وذهبت أرباحها لأولئك الذين طرحوا برامج غير محمية ضد النسخ.

وقد كانت شركة "مايكروسوفت" إحدى الشركات التي أدركت أن التنفيذ بأساليب فنية ليس هو الخيار الوحيد، وكسبت الشركة مكاسب خرافية من جراء بيع منتجات ليست محمية ضد النسخ من الناحية الفنية، وقد اجتهدت الشركة في حث المستفيدين على الالتزام بالسياسات وذلك من خلال وسائل غير فنية. وقد شجعت حوافز التسويق كدعم المستفيدين وتطوير أجهزتهم بتكلفة منخفضة - المستفيدين على شراء تراخيص الاستخدام. وفي هذا السياق كذلك،تم اللجوء إلى الضغوط الاجتماعية لتثقيف المستفيدين، كما لوح باستخدام القنوات القضائية لترهيب من ينتهك سياسات الشركة.

ولا تزال عمليات النسخ غير المرخص بها تكلف منتجي البرمجيات مبالغ طائلة، إلا أن الشركات التي تركز جهودها على إرضاء المستفيدين المسؤولين تستطيع التكيف مع ظروف السوق وتحقيق الانتعاش.

إدارة الوصول على المستودع أو الجهاز الخازن Repository :

تطبق معظم المكتبات الرقمية سياسات على مستوى الجهاز الخازن، أو على مستوى المجموعات. ورغم الاختلاف في التفاصيل، فإن جميع الوسائل تتبع النمط الموضح في الشكل رقم (٧-١). ومن المعروف أن المكتبات الرقمية هي نظم آلية لا مركزية أو موزعة، تتدفق فيها المعلومات من حاسب إلى آخر، وإذا كانت هناك إدارة للتعامل على مستوى الجهاز الخازن، فإنه يمكن إنجاح عملية التعامل هذه محلياً، وما أن تخرج المادة من حدود الجهاز الخازن، حتى تقل فرص التحكم فيها من الناحية الفنية.

وقد سبق لنا تناول قضية الاستخدام اللاحق، وهي تتلخص في أنه ما أن يتلقى الحاسب أي معلومة، حتى يصعب على المدير الأصلي للمكتبة الرقمية أن يسيطر عليها سيطرة فاعلة دون التعرض للمستخدم المرخص له باستخدامها. ومع وجود الشبكات، أخذت المسألة بعداً جديداً، وهو أن النسخ العديدة للمادة الموجودة على الحاسبات المرتبطة بالشبكة بما فيها الذاكرة المخبأة caches والذاكرة المطابقة mirrors والخوادم الأخرى، تكون خارج نطاق سيطرة الجهاز الخازن المحلي.

وقد اكتفت معظم المكتبات الرقمية، وإلى يومنا هذا، بتوفير إدارة الوصول للمعلومات على مستوى الجهاز الخازن مع اعتمادها على الضغوط

الاجتماعية والقانونية لضبط عملية الاستخدام اللاحق، وعادة ما تكون مثل تلك الضغوط في محلها، غير أن بعض الناشرين منزعجون من أن الافتقار إلى السيطرة الكاملة [على مواد المكتبات الرقمية] قد يأتي على عائداتهم المالية، من هنا جاء الاهتمام بالوسائل الفنية التي تضبط عملية النسخ والاستخدام اللاحق، حتى ولو كانت المادة قد خرجت من حدود الجهاز الخازن. وتنقسم هذه الوسائل إلى فئتين، هما: النظم الموثوق بها، والحاويات الآمنة.

– النظم الموثوق بها Trusted Systems :

الجهاز الخازن هو مثال للنظم الموثوق بها، حيث يتوافر لدى مديري المكتبة الرقمية ثقة في أن المكونات المادية للحاسبات الآلية والبرمجيات والإجراءات الإدارية تتمتع جميعها بمستوى ملائم من الأمن والسلامة لاختزان المعلومات القيمة وتوفير ضمانات الوصول إليها. وقد يكون هناك نظم أخرى مرتبطة بهذا الجهاز الخازن وتتمتع بالدرجة نفسها من الثقة. ويمكن للمكتبات الرقمية في داخل شبكة النظم الموثوق بها أن تستخدم وسائل تنفيذ للسياسات تشبه تلك المستخدمة مع الجهاز الخازن الواحد، ويمكن تمرير السمات والسياسات فيما بين النظم مع توفر الثقة من أنها ستلقى معالجة فعالة.

إن تجهيز وعمل شبكات مكونة من نظم موثوق بها ليس بالأمر الهين؛ لإن مكونات النظم الفردية لابد أن تكفل مستوى عالياً من الأمن، وكذلك الحال مع العمليات التي يتم بها تمرير المعلومات فيما بين الحاسبات المختلفة؛ ولهذه الأسباب فإن هذه النظم الموثوق بها تستخدم فقط في مواضع محددة أو في حاسبات مخصصة لأغراض محددة. ولو أن جميع عمليات

تشغيل الحاسبات كانت تتم من المجموعة نفسها أو من قبل مجموعات تعمل في ظل قوانين صارمة، لتضاءلت بالطبع كثير من المشكلات الإدارية. ومن أمثلة النظم الضخمة الموثوق بها شبكة الحاسبات التي تدعم ماكينات الصرف الآلي التابعة للبنوك.

ويمكن القول بأنه لا يمكن التسليم بطبيعة الحاسبات الشخصية للمستفيدين وكيفية إدارتهم لها، ومن المعقول حقاً عدم الثقة بها، ولهذا السبب كان طبيعياً أن تقتصر التطبيقات المبكرة للنظم الموثوق بها في المكتبات الرقمية على الأجهزة المخصصة لأغراض محددة مثل البطاقات الذكية، أو الطابعات الآمنة، أو تقتصر على حاسبات خادمة تستعمل برمجيات شديدة الإحكام.

- **الحاويات الآمنة Secure Containers :**

ما دامت الشبكات غير آمنة، وطالما يصعب تنفيذ النظم الموثوق بها، فإن مجموعات متعددة تعكف على تطوير حاويات آمنة لنقل المعلومات عبر شبكة الإنترنت، حيث يتم من خلالها نقل المادة الرقمية للمستفيد في حزمة تشتمل على البيانات وما وراء البيانات الخاصة بسياسات الوصول للمواد. وعادة ما تكون جميع المعلومات أو بعضها مشفراً. كما أن الوصول للمعلومات يتطلب مفتاحاً رقمياً يمكن تسلمه من نظام دفع آلي أو من أي نظام آخر من نظم التحقق من الهوية. ومن ميزات هذه الطريقة أنها توفر قدراً كبيراً من ضبط عملية الاستخدام اللاحق، كما أن الحزمة المرسلة (البيانات وما وراء البيانات) يمكن نسخها وتوزيعها إلى طرف ثالث، لكن ليس من الممكن الوصول إلى مضمون هذه الحزمة دون ذلكم المفتاح، وتقدم لنا اللوحة رقم (٧-٣) وصفاً

لواحد من هذه النظم.

وبالرغم من ذلك فهناك عائق في تقبل نظم الحاويات الآمنة؛ حيث تنعدم جدواها بالنسبة للمستفيد ما لم يتمكن المستفيد من الحصول على المفاتيح السرية المطلوبة للوصول إليها، ومن ثم الإفادة من محتوياتها. وهو أمر يتطلب نشراً واسعاً لخدمات الأمن ولوسائل الدفع الإلكتروني. وإلى وقت قريب كان انتشار مثل هذه الخدمات محدوداً، لذلك لم يجد الناشرون سوقاً رائجة للمعلومات التي يتم توزيعها عبر تلك الحاويات الآمنة.

اللوحة رقم (٧- ٣)

الشيفرات **Cryptolopes**

ُيستخدم نظام شركة آي بي إم لتشفير الحاويات الآمنة ليتيح للمستفيدين بيع وشراء محتويات المواد من خلال الإنترنت بأسلوب آمن، ويعطينا الرسم الذي يظهر في هذه اللوحة فكرة عن بنية المعلومات في نظام التشفير.

فالمعلومات ترسل في مظروف مشفر يسمى بالحاوية، ويوقع موردو المعلومات على معلوماتهم هذه بالحاوية، ولا يمكن لمن يتلقى هذه المعلومات أن يفتح تلك الحاوية إلا بعد أن يفي بكل المتطلبات التي تفرضها سياسة إدارة الوصول، كأن يدفع مبلغاً مالياً مقابل استخدامه أو إفادته من لتلك المعلومات. ومن الملاحظ أن المحتوى لا يمكن فصله عن عملية إدارة الوصول ومعلومات الدفع في المظروف، أي أنه من الممكن إرسال المظروف لاحقاً إلى الآخرين الذين سيتوجب عليهم الدفع إذا أرادوا فتحه (مع ملاحظة أنه يجب على كل مستفيد أن يحصل على الشيفرة التي تُمكنه من فتح المظروف).

وإضافة إلى المحتوى المشفر، فقد تتضمن الحاويات المشفرة مستخلصاً ذا نص واضح يقدم وصفاً للمستفيدين عن المادة، وتشتمل المعلومات التي يتضمنها هذا المستخلص على ملخص للمادة وتعريفاً بمصدرها ومؤلفها وآخر تحديث لها وحجمها، وسعرها وشروط بيعها. وما أن يقرر المستفيد فتح محتوى هذه الحاوية المشفرة، حتى يصدر له مفتاح رقمي يمكنه من فتحها. ولمشاهدة أحد مفردات هذه المواد بشكل مجاني، ما على المستفيد سوى الضغط على المستخلص، وبالتالي تظهر له المعلومات على سطح المكتب، فإذا أراد مشاهدة المحتوى الذي لابد من دفع مقابل مادي له، على المستفيد الموافقة على شروط استخدام الحاوية المشفرة كما هو موضح في المستخلص.

ويمكن أن تكون محتويات الحاوية المشفرة مركبة أو غير ثابتة dynamic، لأن النظام به إمكانية تغطية خطوط جافا وبرامج جافا وغيرها من البرامج التي تُسكّن المحتوى في حاويات آمنة. وفي سياق الاهتمام بالتوصيف في هذا المجال، رخصت شركة آي بي إم لغة حقوق الملكية الرقمية التي وضعتها شركة زيروكس لتحديد القواعد التي تحكم استخدام المحتوى وتسعيره.

<table>
<tr><td colspan="2">بيان بالمواد</td></tr>
<tr><td colspan="2">نص واضح</td></tr>
<tr><td colspan="2">تعليمات بصمة الأصابع والدفع بالعلامات المائية المشفرة</td></tr>
<tr><td>جزء مشفر من الوثيقة</td><td>تسجيلة أساسية</td></tr>
<tr><td>جزء مشفر من الوثيقة</td><td>تسجيلة أساسية</td></tr>
<tr><td>جزء مشفر من الوثيقة</td><td>تسجيلة أساسية</td></tr>
</table>

الشروط والضوابط
حماية السلامة والتوقيعات

أمن المكتبات الرقمية Security of Digital Library:

نتناول فيما تبقى من هذا الفصل بعض الوسائل الأساسية لأمن المعلومات التي تستخدم في نظم الحاسبات بالشبكة، وهي وسائل لا يقتصر استخدامها على المكتبات الرقمية، بل يتجاوزها لمجالات أخرى، إلا أن المكتبات الرقمية لها حاجات خاصة بسبب أهمية الشبكات اللامركزية المتناهية لكل من موردي المعلومات والمستفيدين منها.

وتبدأ عملية الأمن من إداريي النظم، وهم أولئك الذين يقومون بتجهيز وإدارة الحاسبات الآلية والشبكات التي تترابط فيما بينها. وينبغي أن تكون أمانة هؤلاء فوق كل الشبهات لأن لديهم صلاحية الدخول على بطائن النظام، ويعمل إداريو النظم البارعين على تنظيم الشبكات وملفات النظام حتى يتمكن المستفيدون من الوصول إلى المعلومات المناسبة، كما أن عليهم إدارة كلمات المرور وإعداد ما يعرف "بالجدران النارية firewalls" لعزل أجزاء الشبكات، وتشغيل البرامج التشخيصية تحسباً للمشاكل، وعليهم كذلك عمل نسخ احتياطية من المعلومات حتى يمكن إعادة بناء النظام في حالة حدوث خلل ما، مثل أعطال التجهيزات، أو نشوب حريق، أو حدوث اختراق لإجراءات الأمن.

إن شبكة الإنترنت ليست آمنة؛ إذ يستطيع الكثيرون التلصص ومراقبة المعلومات المتوافرة، وقد يحدث هذا الدخول بشكل مشروع من أجل حل

مشكلة ما، لكنه يمكن أن يحدث لأغراض لا تمت للأمانة والنزاهة بصلة. والسؤال هنا بوجه عام عن الأمن هو:كيف يمكن بناء تطبيقات آمنة عبر هذه الشبكة غير الآمنة ؟

وطالما أن الإنترنت لا تتوافر لها ضمانات الأمن الكافية؛ فإن أمن المكتبات الرقمية يبدأ من الحاسبات التي تشكل بنية هذه المكتبات، ومن المعلومات الموجودة بها، مع تركيز الاهتمام على إجراءات الاتصال بين تلك الحاسبات والشبكات المحلية. وبالنسبة لكثير من الحاسبات الشخصية ليس هناك من سبيل للأمن إلا بتقييد عدد من يستخدمونها، ولبعض الحاسبات الأخرى برامج حماية لا تزيد عن وضع كلمات مرور وأسماء مستخدمين مبسطة، وعندما يشترك أكثر من مستخدم في استعمال الحاسبات يجب اتخاذ إجراءات تحكم لتحديد من يقرأ ومن يكتب في كل ملف.

والخطوة التالية في عملية الحماية تتمثل في ضبط إجراءات الاتصال بين الشبكات المحلية وشبكة الإنترنت الموسعة، وتوفير حاجز في وجه المتطفلين من الخارج، والحاجز الأكمل هو العزل أو الفصل isolation ، أي منع أية اتصالات خارجية بالشبكة. وهناك وسيلة أكثر جدوى في هذا الصدد، وهي أن يتم وصل الشبكة المحلية أو الداخلية بالإنترنت عبر حاسب ذي هدف مخصص يسمى "بالجدار الناري أو جدار الحماية firewall"، الذي يعمل على مراقبة كل حزمة معلومات تحاول المرور عبر الشبكة، ومنع أولئك الذين يحاولون انتهاك إجراءات الأمن. و جدران الحماية هذه يمكنها كذلك منع المحاولات الخارجية للاتصال بالحاسبات

داخل المنظمة أو الشركة، أو أنها تستطيع أن ترفض الحزم التي لا تتفق وقائمة البروتوكولات المعتمدة، ويمكن لهذه الجدر النارية إذا ما أحسن إداراتها أن تكون على درجة عالية من الفعالية في إيقاف المتطفلين عند حدودهم.

ولقد كانت الجامعات في صدارة المؤسسات التي استخدمت شبكات الحاسبات منذ عدة سنوات مضت، ورغم كثرة المستفيدين وتنوعهم، فقد نجحت في وضع نظم آمنة لتأمين شبكات الحاسبات داخل المدن الجامعية campus network التي تضم آلاف الحاسبات، صحيح قد تحدث بعض الانتهاكات وسلوكيات الاستخدام غير السوي في كل جامعة، لكن من النادر أن يحدث عطل كامل في شبكة الحاسبات.

وإذا روعي الحرص والاهتمام في إدارة الشبكات، فإن الحاسبات المرتبطة بإحدى الشبكات يمكن أن يتوافر لها قدر كبير من دواعي الأمن، إلا أن هناك عدة طرق يمكن أن يحتال بها من يريد انتهاك إجراءات الأمن. وفي الجامعات تأتي معظم المشاكل ممن يعملون بها من موظفين ساخطين على أوضاعهم، أو من بعض الطلبة الذين يعملون على سرقة أسماء المستخدمين وكلمات المرور الخاصة ببعض المستفيدين، أما طرق الانتهاك الأكثر فاعلية فهي تتم عن طريق برامج حاسبات خاصة بذلك، ومع أن لكل نظام تشغيل إجراءات الأمن المثبتة معه، فإن المشكلات قد تحدث بسبب الأخطاء في التصميم أو في البرمجة. وهناك برامج هامة جداً للمكتبات الرقمية كالبريد الإلكتروني وخوادم الويب يصعب تأمينها على الوجه الأكمل، ولهذه الأسباب ينبغي على من ينشئ مكتبة رقمية أن يسلم بأن ضمانات الأمن أمر بعيد

المنال، لكن مع الإصرار يمكن تحجيم هذه المخاطر، مع بقاء احتمالات الخطأ. وينبغي على مديري المكتبات الرقمية أن يتخذوا موقفاً متوازناً نحو قضية الأمن، صحيح أن ضمانات الأمن الكاملة مستحيلة، إلا أن تحقيق قدر كبير منها في نظم شبكات الحاسبات دون تكلفة باهظة أمر غير مستحيل، وإن احتاج ذلك بعض الجهد والتركيز.

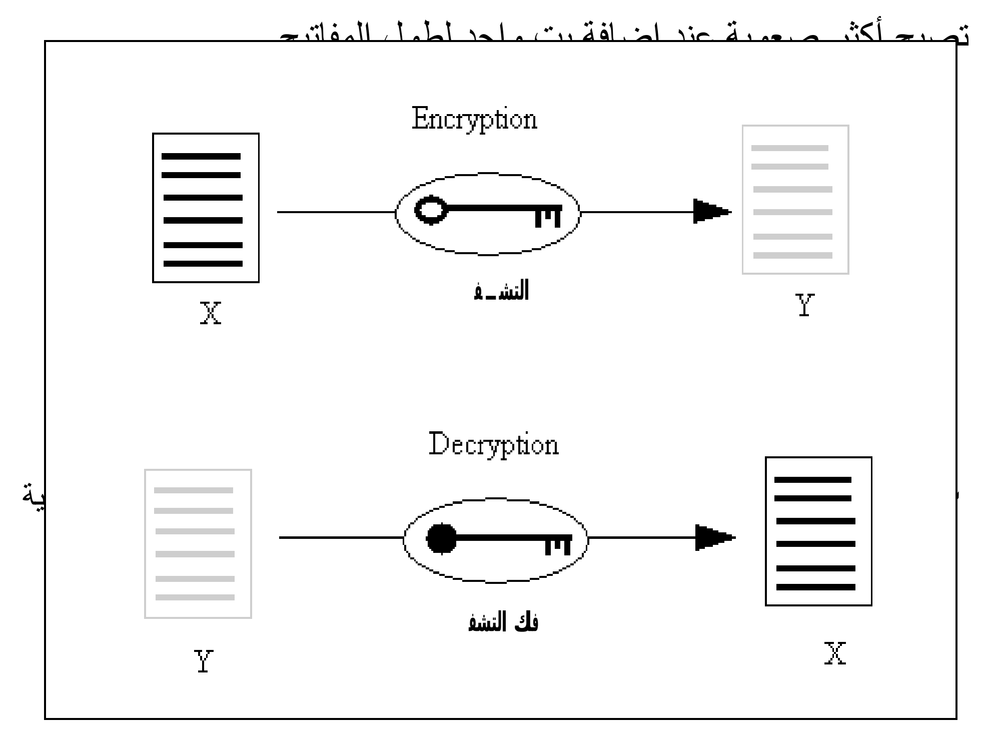
التشفير Encryption :

" التشفير" يقصد به مجموعة الأساليب الفنية المتبعة في تخزين وبث المعلومات الخاصة بطريقة مشفرة تجعلها تبدو في صورة مبعثرة أو غير منتظمة تماماً إلى أن يتم تحويلها عن طريق الإجراءات اللازمة. وحتى لو تمكن شخص غير مصرح له من الاطلاع على هذه المعلومات فلن يتمكن من تغييرها أو العبث بها. ويستخدم التشفير في المكتبات الرقمية في نقل المعلومات السرية عبر الإنترنت، وقد تبلغ بعض المعلومات مستوى من السرية يتحتم معه تشفيرها في كل مكان يتم اختزانها فيه، وتعد كلمات المرور مثالاً واضحاً للمعلومات التي يجب تشفيرها على نحو دائم سواء أكانت مختزنة في الحاسبات أم كانت مبثوثة عبر الإنترنت. وفي كثير من المكتبات الرقمية تعد كلمات المرور هي المعلومات الوحيدة التي يجب تشفيرها.

ويظهر الشكل رقم (٧-٢) الفكرة الأساسية لعملية التشفير، حيث يتم إدخال البيانات المفترض إبقاؤها سرية، والتي يرمز لها بالرمز (X)، في عملية تشفير تنطوي على تحويلات رياضية لتخرج لنا في صورة مشفرة، والتي يرمز لها بالرمز (Y)، وينبغي أن يكون لتلك البيانات التي تم تشفيرها (

Y) العدد نفسه من وحدات البتات التي كانت للبيانات الأصلية غير المشفرة. وقد تبدو هذه البيانات على أنها مجموعة غير منتظمة من الوحدات، لكن يمكن عكس العملية عن طريق مجموعة من الإجراءات المخصصة لذلك بحيث يتم استعادة البيانات في صورتها الأصلية (x)، وكلتا العمليتين - التشفير encryption وفك التشفير decryption- يمكن أن يتم تنفيذهما عن طريق برامج آلية، أو عن طريق أجهزة مخصصة لهذا الغرض.

يتم التحكم في الوسائل شائعة الاستخدام في عملية التشفير عبر زوجين من الأرقام يعرفا "بالمفاتيح keys"، يستخدم أحدهما لأغراض التشفير، ويستخدم الآخر لفك التشفير. وتتنوع وسائل التشفير من حيث الطريقة التي تختار بها المفاتيح. صحيح أن الأشكال الرياضية للعمليات ليست سراً، لكن معايير الأمن تكمن في المفاتيح ذاتها. فالمفتاح عبارة عن سلسلة من البتات التي تتروح ما بين ٤٠ إلى ١٢٠بتاً، لكنها قد تتجاوز هذا الرقم في بعض الأحيان. ومن البدهي أن تكون المفاتيح المطولة أكثر أمناً من المفاتيح القصيرة؛ لأن أي محاولة لانتهاك إجراءات الأمن من خلال تخمين المفاتيح تصبح أكثر صعوبة عند إضافة بت واحد لطول المفاتيح



الشكل رقم (٧-٢)
الفكرة الأساسية لعملية التشفير

وقد اقتصر استخدام التشفير من الناحية التاريخية على قدرات الحاسب؛ فجميع الخطوات تتطلب عمليات حاسوبية كثيرة لمزج المعلومات وإعادة ترتيبها، و قد تطلبت التطبيقات المبكرة لمعيار تشفير المعلومات DES. كما هو موضح في اللوحة رقم (٧-٤) - جهازاً خاصاً يلحق بكل حاسب آلي، أما اليوم وفي ظل وجود الحاسبات ذات السرعات العالية، فلم تعد هذه مشكلة كبيرة، غير أن الوقت المبذول في عملية التشفير وفك التشفير لا يزال أمراً ملحوظاً. وبشكل عام يمكن القول إن أساليب التشفير تثمر عن نتائج جيدة مع الرسائل القصيرة (مثل كلمات المرور)، أو الرسائل بالغة السرية التي تأتي بين حين وآخر، لكنها لا تؤدي النتيجة نفسها مع البيانات الكثيرة، وخاصة عندما يكون لسرعة الرد أهميتها البالغة.

التشفير باستخدام المفتاح الخاص .Private-key Encryption :

التشفير باستخدام المفتاح الخاص هو مجموعة أساليب يستخدم فيها المفتاح لتشفير البيانات وفك تشفيرها مع ضرورة الإبقاء على ذلك سراً. وتعرف هذه العملية أيضاً بالتشفير باستخدام المفتاح الواحد single-key encryption، أو التشفير باستخدام المفتاح السري secret - key encryption. وتقدم اللوحة رقم (٧-٤) وصفاً لمعيار تشفير البيانات الذي يعد من أكثر الأساليب استخداماً.

ولا تبدو جدوى التشفير باستخدام المفتاح الخاص إلا إذا أحاط الأمن بجميع خطواتها، فإذا أراد أحد الحاسبات إرسال معلومة مشفرة لحاسب بعيد، ينبغي عليه أن يجد الطريق الآمن تماماً لتوصيل المفتاح للحاسب

البعيد. لذلك يتم استخدام التشفير بالمفتاح الخاص على نطاق واسع في التطبيقات التي تتولى مهام تبادل المعلومات فيها تلك المرافق أو الخدمات الموثوق بها trusted services.

اللوحة رقم (٧-٤)

معيار تشفير البيانات

The Data Encrytion Standard

معيار تشفير البيانات هو وسيلة تشفير باستخدام المفتاح الخاص، طورته شركة آي بي إم، وأصبح معياراً أمريكياً منذ عام ١٩٧٧م. والواقع أن العمليات الحسابية التي يقوم بتنفيذها هذا المعيار تتسم بالبطء عندما تستخدم في برمجية حاسوبية، لكنها تكون ذات سرعة كافية للعديد من التطبيقات. وحالياً يمكن لأي حاسب شخصي أن يشفر نحو مليون بايت في الثانية الواحدة.

ويستخدم معيار تشفير البيانات مفاتيح مكونة من ٥٦ بتاً. حيث يقوم بتقسيم مجموعة البيانات إلى وحدات، كل منها مكونة من ٦٤بتاً، ويقوم بتشفير كل منها على حدة، ومن المفتاح ذي الـ ٥٦ بتاً يتولد ستة عشر مفتاحاً صغيراً، ويعتمد أساس الخوارزمية على ست عشرة عملية تحويل متتابعة للوحدات ذات الـ٦٤بتاً مستخدماً هذه المفاتيح الصغيرة بالتتابع، ويستخدم التشفير المفاتيح الستة عشر الصغيرة نفسها لتنفيذ التحويلات الراجعة في الاتجاه المعاكس، وقد يبدو هذا وكأنه عملية خوارزمية بسيطة، لكن ثمة فروقاً بسيطة، أهمها أن أنماط البتات المتولدة عن عملية التشفير تبدو عشوائية تماماً وخالية من أية إشارة تدل على البيانات أو على المفتاح.

ويذهب من يغالون في هذه المسألة إلى أن المفاتيح ذات الـ ٥٦ بتاً يمكن كسرها ببساطة عن طريق إجراء عملية التجريب لكل مفتاح يمكن معرفته، لكن هذه مهمة صعبة، والواقع أن معيار تشفير البيانات يناسب تطبيقات المكتبات الرقمية بشكل كاف.

التشفير باستخدام المفتاح المزدوج Dual-key Encryption :

عندما يتم استخدام "التشفير باستخدام المفتاح الخاص" عبر إحدى الشبكات، ينبغي أن يكون هذا المفتاح معلوماً لكل من الحاسب المرسل والحاسب المستقبل. وهو ما يمثل لغز كيفية عمل طرف منهما إذا لم يفلح جهاز الطرف الآخر في إرسال المفتاح إليه بطريقة سرية. ومن هنا جاء " التشفير باستخدام المفتاح المزدوج " ليتيح إرسال جميع المعلومات عبر الشبكة؛ ولهذا السبب اكتسب اسماً آخر، وهو "التشفير باستخدام المفتاح العام public-key encryption"، الذي تظل فيه المعلومات المشفرة حتى لو حدث اعتراض أو رصد لكل رسالة.

وتعد طريقة ريفست - شامير- أولمان Rivest-shamir-Adleman من أشهر طرق التشفير باستخدام المفتاح المزدوج. وهي تتطلب وجود مفتاحين : أحدهما عام والآخر سري، فإذا افترضنا أن شخصاً ما (وليكن (أ)) يريد أن يرسل معلومة مشفرة لشخص آخر (وليكن (ب))، فعليه أن يشفرها مستخدماً المفتاح العام للطرف (ب)، وعندما يتسلم الأخير تلك المعلومة، عليه أن يفك شفرتها مستخدماً المفتاح الخاص والذي لا يعلمه أحد سواه.

يمتاز التشفير باستخدام المفتاح المزدوج هذا بمزايا عديدة، غير أن له عيباً واحداً، وهو أن الأمر يتطلب التأكد من أن المفتاح هو فعلاً المفتاح العام

لشخص بعينه، والحل المعتاد لهذه المشكلة أن تقوم "هيئة تصديق أو إصدار شهادات الترخيص[1] certification authority " موثوق بها بعمل جميع المفاتيح والتثبت من هوياتها أو صحتها، على أن تقوم بعد ذلك بإصدار شهادات certificates عبارة عن رسائل موقعة تحدد هوية كل شخص ومفتاحه العام. وهذه الطريقة لا غبار عليها طالما لم يحدث انتهاك لإجراءات الأمن الخاصة بهيئة إصدار شهادات الترخيص هذه.

التوقيعات الرقمية Digital Signature :

كما سبق أن أسلفنا فإن التوقيع الرقمي يمكن الاستفادة منه في التأكد من عدم العبث بملف معين من ملفات الحاسب، وتعتمد التوقيعات الرقمية على فكرة " دالة المزج أو التمويه، أو القيمة الاختبارية hash function"، وهي دالة رياضية يمكن تطبيقها على بايتات "bytes" أحد ملفات الحاسب لتوليد رقم ثابت الطول.

ومن دوال المزج أو التمويه أو القيم الاختبارية شائعة الاستخدام ما تعرف بـ"إم دي فايف MD5" وهي تقوم بتحويل خاص لبتات bits أحد الملفات لتنتهي إلى ١٢٨بتاً عشوائية تماماً، ويمكن تطبيق هذه الطريقة على ملفات مختلفة الأطوال. وإذا اختلف ملفان ولو في بت واحد، انعكس ذلك الاختلاف على عملية تحويلها، والعكس لو كان هناك ملفان لهما القيمة الاختبارية نفسها hash فإن احتمالات عدم تطابقهما لا يمكن حدوثها تحت أي ظرف. ولذلك فإن من أبسط طرق التأكد من عدم العبث بأحد الملفات هي حساب " القيمة الاختبارية

(١) هناك من يسميها سلطة منح الشهادات الرقمية (المترجمان) .

MD5 hash" الخاصة به عند إنشائه، ثم يعاد بعدها بفترة معينة حساب ذلك ومقارنته بالبيانات الأصلية لرصد أي تغيير طرأ على الملف، فإذا لم تتضح أية اختلافات بين القيمتين، فهذا يعني أن الملف لا يزال كما هو دون حدوث أية تغيرات أو تدخلات.

ولطريقة دالة التحويل إم بي فايف MD5 هذه عدة مزايا منها: أنها الأسرع في حساب الملفات الكبيرة، إلا أنها - شأنها في ذلك شأن أية وسيلة أخرى من وسائل الأمن- تتعرض دائماً لاحتمال قيام مستفيد بارع بكشف طريقة عملها، وإنشاء ملف له قيمة اختبارية hash value محددة. وقد كثر الحديث أثناء إعداد هذا الكتاب عن إمكانية اختراق تلك الطريقة، وإن بقيت لها وظائف أخرى.

ولا تكشف القيمة الاختبارية عن المصدر الذي قام بحسابها، لكن التوقيع الرقمي يخطو خطوة أبعد من ذلك نحو ضمان موثوقية أي كائن بالمكتبة، فعندما يتم حساب القيمة المزجية لكائن ما، فإنها تشفر عن طريق استخدام المفتاح الخاص بصاحب المادة، وهذا ما ينشئ التوقيع الرقمي، إلى جانب المفتاح العام وجهة التصديق. وقبل أن يتأكد المستفيد من القيمة الاختبارية، يقوم بفك شيفرة التوقيع الرقمي مستخدماً المفتاح العام، فإذا تطابقت النتائج دل ذلك أن المادة لم يطرأ عليها أي تغيير، وأن التوقيع الرقمي تم إنشاؤه باستخدام المفتاح الخاص المعني بذلك.

وبرغم كل ذلك، فإن التوقيعات الرقمية ليست مجردة من النقائص، فمع أن المستفيدين من المكتبات الرقمية يرغبون في التأكد من عدم حدوث عبث

في المواد، إلا أنهم لا ينشغلون بدقائق الأمور، بل بمضمون المادة فحسب. فمكتب حق المؤلف في الولايات المتحدة على سبيل المثال يركز اهتمامه الأساسي على المحتوى الفكري، مثل كلمات النص، لكنه لا يعير اهتماماً بما إذا كان الحاسب قد أضاف بعض معلومات الحماية لأحد الملفات أم لا، أو ما إذا كان قد حدث تغيير البنط المستخدم في كتابة النصوص. لكن الفشل الذريع للتوقيع الرقمي يبدو عندما يتغير بت واحد في الملف، وحتى الآن لم تطرح طريقة ناجحة لضمان صحة المضمون وأمنه سوى التأكد من بتاته.

انتشار التشفير باستخدام المفتاح العام :

قد يتوقع المرء أن يكون التشفير باستخدام المفتاح العام قد انتشر منذ سنوات عديدة بحكم أن العمليات الرياضية الأساسية التي اعتمد عليها هذا الأسلوب معروفة منذ عشرين سنة مضت، إلا أن ذلك ليس هو الواقع - مع الأسف الشديد- فلما تزل هناك قضايا فنية هامة يتعلق كثير منها بإدارة المفاتيح وكيفية إنشائها، وكيفية تخزين المفاتيح الخاصة، وما الاحتياطات التي يمكن اتخاذها إذا ما انتهكت إجراءات الأمن في الجهة المسؤولة عن المفاتيح. ومع ذلك ينبغي أن تراعى الأسباب الرئيسية للتأجيل في نشرها بشيء من التصرف الدبلوماسي أو السياسي.

فهناك أولاً مشكلة براءات اختراع البرمجيات (وهو ما سبق تناوله في الفصل السادس). فلا خلاف بين معظم علماء الحاسب الآلي على أن مجال التشفير باستخدام المفتاح العام هو أحد المجالات القليلة في عالم الحاسب الآلي الذي شهد ابتكارات حقيقية. يضاف إلى ذلك أن وسائل التشفير ليست

واضحة، وأن مخترعيها يستحقون المكافأة على اختراعهم، ومع الأسف فإن أصحاب براءات الاختراع ووكلاءهم قد انتهجوا سياسات محدودة في منح التراخيص والتي قلصت من سرعة انتشارها.

لكن العائق الأكبر خطراً تمثل في تدخل الجهات الحكومية الأمريكية، إذ تدعي وكالة الاستخبارات المركزية أن تقنية التشفير تدخل في إطار الأسرار العسكرية، ومن ثم فإن تصديرها للخارج يهدد الأمن القومي للولايات المتحدة الأمريكية، كما يدعي مكتب التحقيقات الفيدرالية أن الأمن العام يعتمد على قدرته في تفسير وقراءة أي رسالة على الإنترنت عند تخويله بذلك. وفيما يخص تصدير التقنيات فإنه من الصعب التسليم بها في ظل الانتشار الواسع لوسائل التشفير في أنحاء العالم، وفي ظل قيام شركات أوربية ويابانية لها وزنها بتسويق منتجات التشفير. أما مسألة الحفاظ على الأمن العام فهي لا تقل تعقيداً، لكن الذي يهون من شأنها أن الشعب الأمريكي لا يثق في القدرات الفنية لوكالة الاستخبارات المركزية وغيرها من هيئات الشرطة ولا في إجراءاتهم الإدارية، ولا هم للناس سوى أنهم يريدون إرسال معلوماتهم السرية دون أن يراقبها أحد.

ومن المناسب أن يختتم هذا الفصل بإشكالية تعكس كيف تقوض مصاعب السياسة قدرات الحلول الفنية. وهو ما يعكس فكرة تتردد في مجال المكتبات الرقمية، ولها أهميتها الخاصة في إدارة الوصول إلى مواد تلك المكتبات، وهي الارتباط الوثيق بين الناس أو المستفيدين من جهة، والتقنية من ناحية ثانية، والإجراءات الإدارية من ناحية ثالثة. والمكتبات الرقمية الناجحة هي التي تجمع بين جوانب جميع هذه الأطراف الثلاثة ولا تعتمد على التقنيات

وحدها.

الفصل الثامن
واجهات المستفيدين ومدى سهولة استخدامها

جرت العادة على تسمية من يستخدم المكتبة "أحد رواد المكتبة patron" أو "القارئ"؛ أما بالنسبة لبيئة الحاسبات فهو يعرف "بالمستفيد User"، أو "المستفيد النهائي end user". وبصرف النظر عن التسمية المختارة فإن المكتبات الرقمية ستفقد قيمتها ما لم يكن من السهل استخدامها بفاعلية.

إن العلاقة بين الإنسان والحاسب الآلي هي مدار أبحاث مكثفة تغطي مجالات متشعبة مثل العلوم المعرفية والتصميم التصويري (الجرافيكي) وعلم البلاغة، والنماذج الرياضية لنظم الحاسبات. وينصرف اهتمام بعض هذه الأبحاث إلى محاولة الوصول إلى فهم نظري لكيفية تفاعل الإنسان مع الحاسب الآلي حتى يتسنى استخدام تلك النماذج الخاصة بمعالجة الإنسان للمعلومات في تصميم نظم حاسوبية مناسبة. وينصرف اهتمام البعض الآخر من هذه الأبحاث إلى مساعدة المستفيد في فهم المبادئ الأساسية التي تكمن وراء النظم الحاسوبية، ومن ثم تحفيزه على الإفادة الفعالة من الخدمات والمعلومات التي تتيحها تلك النظم. وستقتصر طموحاتنا في هذا الفصل على عرض الوسائل شائعة الاستخدام في هذه الأيام، مستعرضاً بعض شواهد النظم التي لا تزال في طور التجربة، والتي تبدو واعدة على كل حال.

إن التغير هو إحدى قضايا المكتبات الرقمية، كما يعد في الوقت ذاته إحدى المشكلات التي تواجه تصميم واجهات المستفيدين. وإذا كان استخدام

المكتبات التقليدية بصورة فعالة أمراً صعباً، فإن معدل التغير فيها يتسم بالبطء مما يتيح لروادها أن يتمرسوا فيه ويعتادوا عليه، بل يصبحوا مع مرور الأيام خبراء في هذا المجال، أما بالنسبة للمكتبات الرقمية فمعدل التغير والتطور فيها يبلغ من السرعة ما يصل إلى إضافة خدمات جديدة، ومجموعات جديدة، وواجهات مستفيدين جديدة في كل شهر، وهو ما يعني مزيداً من الإرهاق الذهني لمتابعة مثل هذه التطورات. وإذا كان من عادة المستفيدين أنهم لا يأنسون إلى ضرورة تعلمهم مهارات أساسية من جديد، لكن يبدو أن التغيير شر لابد منه، إن جاز التعبير.

ويعود شيء من التباين الشديد في مستويات خبرة المستفيدين إلى السرعة التي يحدث بها التغير في المكتبات الرقمية. وقد انبثقت معظم آفاق التطور في المكتبات الرقمية من محيط الجامعات حيث يتواجد بها كثير من الخبراء. وحيث يجد المستفيدون يد الزملاء والمكتبيين ممدودة عوناً لهم، وحيث يتفرغ إداريو النظم هناك لتهيئة الحاسبات وتحميل البرامج عليها، وتتبع ما يطرأ على السوق من تقلبات. ولكن مع انتشار الإنترنت، بدأت المكتبات الرقمية تُستخدم من قبل أناس ليس لهم معرفة أو صلة بذوي الخبرات، كما أنهم لا يحبذون تضييع أوقاتهم في تعلم أشياء يظنون أن ليس لها سوى قيمة مؤقتة، ولعل ذلك ما يسبب شيئاً من التوتر؛ فالإمكانات المتقدمة [للحاسبات] تعد قيمة بالنسبة للمتخصصين، كما أنها تمكن المستفيدين المهرة من تحقيق معدلات عمل أسرع وبفاعلية أكبر، أما المكتبات الرقمية فيجب أن تُهيأ لمن تتوافر لهم الحدود الدنيا من التدريب

على استخدامها.

مدى الاستخدام وتصميم واجهات المستفيدين :

من السهل عند مناقشة إمكانية استخدام أي من نظم الحاسبات الآلية، التركيز على تصميم الواجهة بين المستفيد والحاسب، غير أن سهولة الاستخدام هي خاصية تحسب للنظام كله، وينبغي أن تعمل جميع مكونات المكتبة الرقمية بانسياب معاً لجعل المكتبة أكثر فاعلية في أداء دورها، وجعلها في متناول كل من الرواد، والمكتبيين، وإداريي النظم.

ويوضح الشكل التالي رقم (٨-١) تصور العلاقة بين مدى الاستخدام، وتصميم واجهات المستفيدين. مع ملاحظة أن واجهات المستفيدين لنظم الحاسبات تقوم على "نموذج تصوري conceptual model" يصف الطريقة التي يستخدم بها النظام. ونورد هنا بعض النماذج التصورية المستخدمة في تصميم المكتبات الرقمية (وميدانياً يلاحظ أن معظم المكتبات الرقمية تجمع بين تصورات مختلفة من هذه النماذج).

- في نموذج المكتبة القديمة: يقوم المستفيد بالبحث في الفهرس أو الكشاف، ثم يختار بعض الكائنات مما يظهر له من نتائج البحث، ثم يذهب بعد ذلك لاسترجاع هذه الكائنات من المستودعات.

- في النموذج الأساسي للويب: يقوم المستفيد بتتبع الروابط الفائقة التي تربط بين الملفات.

- يقوم البروتوكول Protocol Z39.50 بدعم نموذج تصوري للبحث في إحدى المجموعات ثم اختزان مجموعة من النتائج لأغراض المعالجة

والاسترجاع اللاحق.

الشكل رقم (٨ – ١)

العلاقة بين مدى الاستخدام وتصميم واجهات المستفيدين

النموذج التصوري Conceptual Model	تصميم الواجهة Interface Design
	التصميم الوظيفي Functional Design
	البيانات،وما وراء البيانات Data & Metadata
	نظم الحاسبات والشبكات Computer Systems and Networks

يبين الجانب الأيسر من هذا الشكل (٨-١) مستويات التصميم المطلوبة لتنفيذ أي نموذج تصوري، ويأتي في أعلاها تصميم واجهة المستفيدين، ويشمل هذا المستوى كل ما يظهر على الشاشة وكيف يقوم المستفيد بالتعامل معه، ومن أمثلة ذلك: أشكال الخطوط، والألوان، والشعارات، ومراقبات لوحة المفاتيح، وقوائم التشغيل والأزرار. ثم يأتي بعد ذلك "التصميم الوظيفي" ويتم فيه تحديد الوظائف المتاحة للمستفيد، وتشتمل هذه الوظائف على اختيار أجزاء من إحدى الكائنات الرقمية، والبحث في إحدى القوائم أو فرز النتائج، والحصول على المساعدة، والتعامل مع الكائنات الرقمية التي تظهر على الشاشة. ومن الملاحظ أن هذه الوظائف تصبح قابلة للتغيير من خلال "البيانات وما وراء البيانات" التي توفرها المكتبة الرقمية، كذلك عن طريق "نظم الحاسبات والشبكات". وتقدم اللوحة رقم (٨-١) تفاصيل عن

كيفية تطبيق هذه الجوانب لعملية مدى الاستخدام على أحد التطبيقات الهامة.

إن الغالبية العظمى من الحاسبات الشخصية الحالية تقريباً تتوافر لها واجهات المستفيدين على غرار الأسلوب الشائع في حاسبات "آبل ماكنتوش" وهو الأسلوب المستمد من البحوث المبكرة لشركة زيروكس بارك Xerox PARC. وهذا الأسلوب يستخدم الصورة أو أيقونات metaphor الملفات والمجلدات على سطح المكتب. وتشتمل خصائصها على النوافذ المتداخلة، والقوائم، وأدوات الإشارة كالفأرة. ورغم المحاولات المتعددة لتحسينه، لا يزال هذا الأسلوب هو المسيطر على سوق الحاسبات. لدرجة أنه لو قُدّم لأحد مستخدمي حاسبات تعمل على بيئة ماكنتوش منذ خمسين سنة، حاسب يعمل بأحدث نظم مايكروسوفت، لوجد أن واجهة تعامل المستفيدين مألوفة له. صحيح أنه قد حدثت بعض الابتكارات الجديدة، كما تطورت المكونات أو التجهيزات المادية تطوراً كبيراً، إلا أن الصورة الأساسية لا تزال هي نفسها المستخدمة.

إن النموذج التصوري هو صورة سطح المكتب، حيث تبدو فيه الملفات وكأنها مستندات يمكن نقلها إلى سطح المكتب، أو وضعها في مجلدات، أو تخزينها على أقراص. ولجميع واجهات المستفيدين المعتمدة على أسلوب النوافذ النموذج التصوري نفسه. فعلى سبيل المثال يستخدم نظام آبل الفأرة ذات الزر الواحد، أما نظام مايكروسوفت فيستخدم الفأرة ذات الزرّين، وعادة ما يستخدم نظام يونكس الفأرة ذات الأزرار الثلاثة. وتشتمل الوظائف التي تدعم هذا النموذج على عمليات فتح الملفات والمجلدات وإغلاقها واختيارها ونقلها من مكان لآخر، إلى آخر هذه العمليات. ومن الملاحظ أن

الوظائف لا تختلف كثيراً من إحدى الشركة المصنعة إلى أخرى، لكن الاختلاف بين النظم يكمن في "ما وراء البيانات" التي تستخدم لدعم هذه الوظائف. فصورة سطح المكتب تحتاج إلى التطبيقات التي تتعامل مع ملفات البيانات. في حين تستخدم نظم مايكروسوفت ويونكس طريقة تسمية الملفات file-naming، حيث تنتهي على سبيل المثال أسماء الملفات التي تستخدم مع عارض الملفات في صيغة الوثائق المحمولة المعروفة بي دي إف PDF Viewer بالامتداد pdf، أما نظام آبل فيقوم باختزان ما وراء البيانات هذه في هياكل بيانات منفصلة، ومختفية عن المستفيد. وتسمح هذه الاختلافات في نظم الحاسبات الآلية التي تتعامل مع المستفيدين لبعض واجهات المستفيدين بأداء كثير من المهام في آن واحد.

اللوحة رقم (٨ - ١)

تقليب الصفحات **Page turning**

يبين تطبيق تقليب الصفحات جوانب تصميم "واجهة المستفيدين" التي تظهر في الشكل رقم ٨ - ١). ولكل مشروع من مشروعات التحويل، مثل جستور JSTOR، والذاكرة الأمريكية، مجموعاته من الكائنات الرقمية، وتتكون كل منها من صور لصفحات تم أخذها عن طريق الماسح الضوئي لكتاب أو غيره من المواد المطبوعة الأخرى. وتقوم فكرة "النموذج التصوري" هنا على فكرة تعامل المستفيد مع أي كائن [رقمي] كتعامله تماماً مع الكتاب. والتي تتمثل عادة في القراءة المتتابعة لصفحات الكتاب على الترتيب، غير أن القارئ يستطيع كذلك أن يرجع إلى صفحة سابقة، أو ينتقل إلى أي صفحة أخرى. ومن ثم تبرز الأهمية خاصة لبعض الصفحات

في مثل هذه الحالات كصفحة قائمة المحتويات أو الكشاف. وحيث إن هناك كثيراً من الحاسبات الشخصية التي تكون شاشاتها أقل من حجم الصفحات المطبوعة، وذات درجة وضوح أقل من تلك التي تمتع بها صفحات الكتاب، فإن النموذج التصوري يقدم خاصيتي تقريب زاوية الكاميرا وتبعيدها zooming وخاصية تحركيها panning عبر الصفحة الواحدة.

يعمل أسلوب تصميم "واجهة المستفيدين" على تحديد المظهر الفعلي على الشاشة، مثل اختيار الإطارات، والأيقونات، والألوان، والمفاتيح البصرية visual clues. كما يشتمل أيضاً على القرارات المتصلة بكيفية تقديم الوظائف الفردية للمستفيد، كما يعمل تصميم "واجهة المستفيدين" كذلك على تعيين الشكل الذي تظهر به الأيقونات، وما يكتب على الأزرار، وتحديد أماكن وجودها على الشاشة. كما يحدد ما إذا كانت عملية التقريب أو التبعيد والتحريك تتم بشكل مستمر أم تخضع لخطوات دقيقة. وتجدر الإشارة إلى أننا قمنا بتصميم "مقلب الصفحات page turner" في جامعة كارنيجي ميلون، وصل تصميم "واجهة المستفيدين" بمساحة الشاشة المخصصة لعرض صور الصفحات إلى غايتها القصوى، وكان التحكم في معظم عمليات المعالجة يتم عن طريق لوحة المفاتيح؛ حيث كان استخدام مفاتيح الأسهم arrow keys يتم لأغراض التحرك حول صفحة بعينها، في حين كان مفتاح المسافات يستخدم لأغراض الانتقال للصفحة التالية. وكان من الممكن لنا استبدال هذا التصميم بتصميم آخر يتم فيه التحكم في تلك الوظائف عن طريق الأزرار الموجودة على الشاشة، لكن رأينا أن ذلك سيكون على حساب المساحة التي ستخصص لعرض صور الصفحات على الشاشة.

أما التصميم الوظيفي فينبغي له أن يوفر الخدمات المناظرة لوظيفة تقليب صفحات الكتب ذاتها، بما ذلك الانتقال إلى الصفحة الأولى أو التالية أو السابقة أو الأخيرة، كما ستكون هناك وظائف تتصل بالتعامل مع مضمون صفحات بعينها مثل الانتقال إلى الصفحة التي تحمل رقم كذا، أو الانتقال إلى صفحة قائمة محتويات الكتاب. كما تساعد الوظائف الأخرى لواجهة المستفيدين في عملية التقريب أو التبعيد أو التحريك داخل أي صفحة عن طريق تحريك المنطقة المعروضة إلى أعلى الشاشة أو أسفلها وتكبير هذه المنطقة المعروضة أو تصغيرها.

وتعتمد الوظائف التي توفرها "واجهة المستفيدين" على نوعية الكائنات الرقمية في المجموعات، وخاصة على ما وراء البيانات البنائية structured metadata. والواقع أن صور الصفحات هي صور يتم اختزانها كملفات مضغوطة يمكن استرجاعها في أي تسلسل. ولكي يتم تقليب الصفحات بشكل متتابع، ينبغي لما وراء البيانات البنائية أن تحدد صورة الصفحة الأولى وعمل قائمة بتسلسل الصور الباقية الأخرى وتحتاج عملية الانتقال إلى إحدى الصفحات المرقمة إلى ما وراء بيانات بنائية تقوم بربط أرقام الصفحات بالتتابع التي تكون عليه الصفحات، لأنه من النادر أن تكون الصفحة الأولى لإحدى المجموعات هي الصفحة التي تحمل رقم (١). كما تحتاج وظيفة تقريب الصورة أو تبعيدها أو تحريكها إلى ما وراء بيانات تحدد أبعاد كل صفحة.

من الملاحظ أن فاعلية أداء واجهة المستفيدين تتوقف على أداء النظام

الآلـي الذي تنفذ عليه، فقد يسبب تأخر إرسال صورة إحدى الصفحات عبر إحدى الشبكات إزعاجاً للمستفيد، وأحد الحلول التي يمكن طرحها لتخفيف هذا الإزعاج هو توقع حاجات المستفيد عن طريق إرسال الصفحات من المستودع إلى حاسب المستفيد قبل أن يطلبها المستفيد، فتكون الصفحة التالية على الأقل جاهزة في الذاكرة، وهذا ما يعرف بالتجهيز أو التفتيش المسبق. pre-fetching وفي المثال الذي ضربناه لمقلب الصفحات في مشروع جامعة كارنيجي ميلون كانت الأولوية لتحقيق الاستجابة السريعة، والتي كانت تقدر بثانية تقريباً لقراءة صورة الصفحة وعرضها ثم بثها عبر الشبكة المزدحمة للجامعة. وقد أدى ذلك إلى التنفيذ المتداخل pipelined implementation(١) الذي ينقل فيه الجزء الأول من الصفحة معروضاً على حاسب المستفيد قبل أن يكون الجزء الأخير منها قد قرئ في الحاسب المستودع من الأسطوانة المخزنة عليها.

المتصفحات browser :

لقد كان لدخول المتصفحات - وخاصة متصفح الموزايك mosaic الذي ظهر في عام ١٩٩٣م - دور كبير في التشجيع على تصميم واجهات تعامل أفضل لأغراض التطبيقات الشبكية. وبالرغم من أن المتصفحات قد صممت لاستخدامات الويب، فإن المرونة التي تتمتع بها جعلت من الممكن استخدامها كواجهات تعامل لمعظم أنواع التطبيقات على الإنترنت، بما في ذلك المكتبات الرقمية. وقبل ظهور المتصفحات ذات الأغراض العامة

(١) يعرف أحياناً بالتنفيذ على خط التعليمات، أو التنفيذ بأسلوب خطوط الأنابيب، وهو يعني تشغيل سلسلة من التعليمات قبل اكتمال سلسلة أخرى (المترجمان).

general-purpose browser، كان على مصممي البرامج تقديم واجهة تعامل مستقلة لكل نوع من أنواع الحاسبات، ولكل بيئة حاسوبية. وكان من الضروري تعديل أي من هذه الواجهات كلما تغيرت نظم التشغيل الآلية. وهي مهمة صعبة لم يقدم عليها إلا قلة، بل لم يفلح أي منهم في القيام بها على أكمل وجه. وبالاعتماد على متصفحات الويب في التفاعل الفعلي مع المستفيد، يستطيع مصمم المكتبة الرقمية الآن أن يركز على كيفية تنظيم عملية تدفق المعلومات إلى المستفيد، وترك مهام التعامل مع تعقيدات المكونات المادية ونظم التشغيل للمتصفح.

الوظائف الأساسية:

تتركز الوظيفة الأساسية لأي متصفح على استرجاع ملف بعيد من جانب خادم للويب ونقلها إلى حاسب المستفيد.

- ولتحديد مكان الملف على خادم الويب، يحتاج المتصفح إلى محدد موحد لموقع المصدر URL. وهو ما يستطيع المستفيد كتابة اسمه على الشاشة، أو أن يكون له رابط من داخل صفحة مهيأة بلغة تهيئة النصوص الفائقة HTML، أو يمكن تخزينه كإشارة مرجعية أو علامة bookmark [(١)]، في الملف الخاص بذلك.

- ومن المحدد الموحد لموقع المصدر (URL) يقوم المتصفح باستخلاص البروتوكول من هناك، فإذا كان البروتوكول هو بروتوكول نقل الملفات

(١) تختلف تسمية هذه العملية بين المتصفحات حيث تسمى bookmark في Netscape navigator وfavorite في Internet explorer و Hotlist في الموزايك (المترجمان).

الفائقة HTTP يقوم المتصفح باستخلاص اسم مجال أو نطاق للحاسب الذي خزن عليه الملف من المحدد الموحد لموقع المصدر (URL). بعد ذلك يقوم المتصفح بإرسال رسالة واحدة عبر بروتوكول نقل الملفات الفائقة (HTTP)، وينتظر الاستجابة ثم يغلق الاتصال.

- وإذا سارت الأمور بشكل جيد، تكون الاستجابة مكونة من ملف، ونوع المادة التي يحملها الملف MIME Type. ولنقل الملف إلى حاسب المستفيد، يقوم المتصفح بفحص نوع هذه المادة واتخاذ الإجراءات الملائمة، وقد تكون هذه الإجراءات مبيتة داخل المتصفح، أو قد تكون عبارة عن برنامج خارجي يستدعيه المتصفح. لن يستطيع أي متصفح من متصفحات الويب البسيطة أن يقدم أكثر من المساندة أو الدعم لبروتوكول نقل الملفات الفائقة، واتخاذ إجراءات عرض الصفحات في صيغة تهيئة النصوص الفائقة، وبحكم وضوح الرؤية والدقة التي يتمتع بها كل من بروتوكول نقل الملفات الفائقة (HTTP) ولغة تهيئة النصوص الفائقة (HTML)، يجب أن يتحرر برنامج المتصفح من أية تعقيدات.

توسعة مجال عمل المتصفحات إلى خارج نطاق الويب :

بالرغم من أن المتصفحات طورت لأغراض خدمة الويب، وأن كل متصفح يعمل على دعم البروتوكولات الأساسية للويب وقليل من الصيغ القياسية؛ فإن عمل هذه المتصفحات يمكن أن يتسع ليشمل تقديم خدمات أخرى أثناء تأديته وظائفه الأساسية كواجهة تفاعل للمتصفح. ويعود نجاح المتصفحات والويب، بل والإنترنت جميعها إلى هذه المرونة في توسعة مجال العمل.

ويتمتع المتصفح المعروف بالموزايك وجميع أجياله بالأشكال الثلاثة نفسها من المرونة في التوسعة والتي تتمثل في مرونته في التعامل مع جميع أنواع البيانات، ومع البروتوكولات، وفي تنفيذ البرامج.

أنواع البيانات Data Type :

مع كل نوع من أنواع البيانات تقوم المتصفحات ببعض الإجراءات الروتينية لإحضار ملفات ذلك النوع. وقد تم بناء عدد قليل من الأنواع في جميع المتصفحات الحديثة، بما في ذلك نوع النص البسيط Plain Text، وصفحات النصوص المهيئة بلغة تهيئة النصوص الفائقة (HTML)، والصور في صيغة ملفات الرسومات التصويرية (GIF)، وفي مقدور المستفيدين إضافة أنواع أخرى من خلال بعض الآليات مثل تطبيقات المساعدة helper، والتوصيلات plug-ins. ويذكر أن برنامج تطبيقات المساعدة هو برنامج مستقل يتم تشغيله عن طريق أنواع منتقاة من البيانات، حيث يرسل الملف المصدر إلى المساعد على أنه مجموعة بيانات. وعلى سبيل المثال، عندما لا يكون لدى المتصفحات أساليب دعم داخلي للتعامل مع ملفات في صيغة بوست سكريبت Postscript لكن يتوافر لكثير من المستفيدين عارض ملفات بوست سكريبت في حاسباتهم والذي يستخدم كتطبيق المساعدة. ومن ثم فعندما يتلقى المتصفح ملفاً من نوع بوست سكريبت هذا يقوم بتشغيل برنامج العرض ويرسل إليه الملف لكي يعرضه. أما تطبيق الوصل plug-ins فيشبه تطبيق المساعد، فيما عدا أنه ليس برنامجاً مستقلاً، وإنما يستخدم لأغراض عرض مصدر الملفات للصيغ غير القياسية، مع الملفات المكتوبة بصيغة

لغة تهيئة النصوص الفائقة، وفي عرض واحد.

البروتوكولات :

يعد بروتوكول نقل الملفات الفائقة المعروف بـ HTTP البروتوكول الرئيسي للويب، مع أن المتصفحات تدعم البروتوكولات الأخرى كذلك، والتي كان لبعضها مثل جوفر Gopher وويز WAIS، الأهمية التاريخية ذاتها؛ لأنها سمحت للمتصفحات الوصول إلى خدمات معلومات أقدم. أما غير ذلك من البروتوكولات مثل، نت نيوز Netnews والبريد الإلكتروني، وبروتوكول نقل الملفات FTP، فلا تزال لها أهميتها. ومن نقاط الضعف التي تحيط بمعظم المتصفحات أن قائمة البروتوكولات التي تدعمها عادة ما تكون محددة وغير قابلة للتوسعة. ولذلك ليس هناك من وسيلة معتادة لإضافة البروتوكول Z39.50، أو غيره من البروتوكولات الأخرى إلى المتصفحات.

تنفيذ البرامج :

إن إحدى رسائل بروتوكول نقل الملفات الفائقة (HTTP) التي يرسلها من أحد المتصفحات هي أكثر من مجرد استرجاع ملف معلومات ساكن Static من أحد الحاسبات الخادمة. وإنما يمكنها أن تعمل على تشغيل أحد البرامج على الخادم، ومن ثم تعيد النتائج إلى المتصفح. وكان الأسلوب المبكر لتحقيق ذلك يتمثل في واجهة بوابة العبور العامة Common gateway interface (CGI)، التي تستطيع أن توفر طريقة مبسطة للمتصفح ليتمكن من تنفيذ برنامج معين على حاسب بعيد. وغالباً ما تعرف برامج بوابة العبور

العامـة بخطـوط أو بنـصوص واجهـة بـوابة العـبور العامـة CGI script [١].
وواجهة بوابة العبور العامة هذه هي الآلية التي تستخدمها معظم برامج بحث الـويب لإرسـال التـساؤلات مـن المتـصفح إلـى نظـام الـبحث. وعلـيه يقـوم الناشـرون باختـزان مجمـوعاتهم فـي قـواعد البـيانات ويـستخدمون نـصوص واجهـة بـوابة العبور العامـة (GCI) في تمكين المستفيدين من الوصـول إلـى تلك المجموعات. ومن ثم فـإن التفسير غيـر الرسـمي للمحدد الموحد لموقع المصدر المكتوب هو:

htttp://www.dlib.org/cgi-bin/seek?author = ” Arms ”

يعنــي: علــى الحاســب الـذي يحمـل اسـم الـنطاق www.dlib.org تنفـيذ البـرنامج فـي الملف المسمى cgi-bin/seek، قم بالبحث عن المتغير المسمى " Arms" ، ثـم قم بعـد ذلـك باسـترجاع نتـيجة هـذه العملـية؛ بـناءً علـى ذلـك فـإن البـرنامج يقـوم بـالبحث فـي قاعدة البيانات عن التسجيلات التـي وردت فـيها الكلمة Arms في حقل المؤلف تحديداً.

وقد كانت الاستخدامات المبكرة لواجهة بوابة العبور العامة منصبة على وصـل المتـصفحات بقـواعد البيانات الأقدم وغيرها من المعلومات الأخرى، ولكـن الآن وفـي تطـور غـريب غيـر متـوقع وبعـد أن أصـبحت الـويب نظامـاً ناضـجاً، انعكست الأدوار؛ حيث يقـوم مطورو المكتبات الرقمية المتقدمة في الغالـب باسـتخدام واجهـة بـوابة العـبور العامـة كطـريقة لـربط الـنظام القـديم (الويب) بأنظمته الآلية الأحدث.

الرمز المتنقل Mobile code :

(١) النصوص المحتوية على أوامر برنامج صغير (المترجمان).

لقد أصبحت المتصفحات وبخطى سريعة مجالاً تجارياً كبيراً بفضل المزايا الحديثة التي تضاف على نحو مستمر، مع ملاحظة أن بعض هذه المزايا جاءت في صالح المستفيدين، في حين جاء بعضها الآخر لصالح المنتجين. وقد شملت التطورات كثيراً من الجوانب، مثل تحسين الأداء، وتطوير لغة تهيئة النصوص (HTML)، وأدوات الدعم المبينة داخلياً للصيغ الأخرى، تمثلت الإضافات المبكرة في تخزين ملفات الصور وفقاً لمعيار (jpeg)، وتمثلت كذلك في تطوير طرق أفضل لإضافة صيغ جديدة. ولعل أهم تغيرين حظيا بأهم التحسينات تمثلا في : الإضافة الثابتة لخصائص إجراءات الأمن، وتقديم الرمز المتنقل (الذي يسمح للحاسبات الخادمة بإرسال برامج آلية إلى الحاسب العميل ليقوم المتصفح بتنفيذها على حاسب المستفيد.

ويمنح الرمز المتنقل Mobile Code لمصمم موقع الويب القدرة على إنشاء صفحات الويب التي تضم برامج حاسوبية، وتصف اللوحة رقم (٨-٢) إحدى هذه الطرق التي تستخدم برامج الأبليت Applets [١]، المكتوبة بلغة جافا للبرمجة. وبرنامج الأبليت يمكن نسخه من موقع الويب إلى برنامج العميل، ويتم تنفيذه على الحاسب العميل. ولأن لغة جافا لغة برمجة مكتملة المعالم، فيمكن استخدامها للعمليات المعقدة مثل إرسال نموذج ضمانات الثقة من موقع الويب إلى أحد المتصفحات، ومن ثم يمكن للمستفيد كتابة بيانات التحقق من هويته كاسم المستخدم وكلمة السر التي يقوم برنامج الأبليت بتشفيرها ثم إعادة إرسالها مرة أخرى بأمان إلى الحاسب الخادم.

وليست لغة جافا للبرمجة هي السبل الوحيدة لتوفير الرمز المتنقل، وإنما

(١) برامج آلية صغيرة (المترجمان).

ثمة بديل آخر يتمثل في أن تتضمن الصفحة المهيأة بلغة تهيئة النصوص الفائقة نصوصاً من التعليمات، عادة ما تكون مكتوبة بلغة جافا سكريبت Java script، وهي لغة أسهل من لغة جافا للبرمجة، سواء من حيث كتابتها وسرعة تنفيذها. وتتمثل أبسط استخدامات جافا سكريبت في فحص البيانات التي يقدمها المستفيد كمدخلات بمجرد كتابتها، وبدون تأجيل إرسال أي شيء مرة أخرى إلى الحاسب الخادم للتأكد من صحته. ننوه هنا إلى أن لغة جافا للبرمجة، ولغة جافا سكريبت لغتان مختلفتان تماماً، وما جاء التشابه في الاسم إلا لأغراض تجارية فقط. وبالرغم مما حظيت به لغة جافا للبرمجة من شهرة أكبر، فإن لغة جافا سكربيت تتمتع هي الأخرى بمزايا كثيرة، وعلى أية حال فكلتاهما تستخدمان على نطاق واسع.

اللوحة رقم (٨ – ٢)

لغة جافا للبرمجة **Java**

جافا هي لغة برمجة ذات أغراض عامة، وقد صممت أصلاً لأغراض إنشاء نظم موزعة لا مركزية، وخاصة نظم واجهات المستفيدين التي تعمل في بيئة شبكية.

لقد كان السبيل العادي لعمل برنامج واجهة تفاعل تعمل على عدة أنواع من الحاسبات (مثل متصفحات الويب، ونظم البريد الإلكتروني) يتمثل في كتابة عدة إصدارات [من هذا البرنامج]. وحتى عند كتابة هذه النسخ بلغة معيارية واحدة، مثل لغة (C)، فإن الاختلافات بين بيئة النوافذ واليونكس، والماكينتوش، كانت تجبر المصمم على كتابة عدة إصدارات، وتعديلها باستمرار كلما تغيرت نظم التشغيل. كذلك كان على المستفيد

الـذي يـرغب فـي تـشغيل واجهـة جديـدة أن يقـوم أولاً بإيجـاد إصـدارة [برنامج] هذه البرمجية ثم تركيبها حسب نوع حاسبه. وقد كان هذا - في أحسن الظروف - يؤدي إلى إهدار الوقت وإضاعته، كما أن هذا البرنامج الجديد كان من الممكن - في أسوأ الظروف- أن يؤدي إلى تعطيل تشغيل بعض البرامج الموجودة، أو إحداث بعض الفيروسات.

إن البرامج الحاسوبية عادة ما تكتب بلغة عالية المستوى تعرف بالرمز المصدري source code [١]، وبعد كتابة البرنامج، عادة ما يتم تجميعه في لغة الآلة للحاسب المحدد. وجامع لغة الجافا JAVA Compiler مختلف على أية حال، فبدلاً من إنشاء رمز الآلة لنظام آلي محدد، فإنه يقوم بتحويل مصدر لغة الجافا إلى رمز وسيط يعرف بجافا بايتكود java bytecode ، هذا الرمز موجه لبيئة برامج تسمى آلة جافا التخيلية Java virtual machine. ولتشغيل البايتكود على حاسب معين فإن ذلك يتطلب خطوة أخرى، يتم فيها ترجمة كل جملة مكتوبة بالبايتكود إلى تعليمات رمز الآلة machine-code instruction لهذا الحاسب، ثم يتم تنفيذها. وتعمل المتصفحات الحالية على دعم آلة جافا التخيلية، مع تضمينها برامج ترجمة الجافا Java interpreter.

أما جافا أبليت فهو برنامج آلي صغير يجمَّع في ملف من ملفات جافا بايتكود، ويمكن إرساله عبر الشبكة إلى المتصفح عن طريق أوامر بروتوكول نقل الملف الفائق HTTP. ويقوم المتصفح بالتعرف إلى الملف كبرنامج أبليت، ومن ثم يطلب من مترجم جافا أن ينفذ ما به من تعليمات.

---

(١) الشيفرة المصدرية (المترجمان).

وبما أن جافا لغة برمجة مكتملة المعالم فيمكن بالتالي إدخال معظم الإجراءات الحاسوبية أو الآلية في تطبيقات الويب. كذلك يمكن لنظام جافا أن يمد المبرمجين بمجموعة من الأدوات التي يمكن تضمينها مع البرامج، وتشمل هذه الأدوات المكونات الأساسية للبرمجة، مثل مجموعة الرموز، أو الأرقام، والمدخلات والمخرجات، وهياكل البيانات، كذلك يمكن أن يقدم الأساليب التي يمكن استخدامها في بناء برامج الأبليت، بالإضافة إلى خدمات الشبكة، مثل المحددات الموحدة لمواقع المصادر URLs، ومقاييس بروتوكول التحكم بالإرسال TCP sockets، وعناوين بروتوكول الإنترنت IP address، كما أنه يساعد على تقديم المساعدة في كتابة البرامج التي يمكن إعدادها للنصوص واللغات غير الإنجليزية، ناهيك عن توفير ضمانات الأمن التي تحتوي على التوقيع الإلكتروني، وبرامج إدارة المفتاح العام أو الخاص، والتحكم في الوصول، والتثبت من الهوية، كذلك يقدم المكونات البرمجية المعروفة بجافا بينزJava beans والتي يمكن تركيبها مع بنيات برمجيات أخرى، وأخيراً دعم إمكانات الاتصال بقواعد البيانات. وبما أن الوظائف الأساسية تعد مكونات دائمة لمتصفح الويب، فليست هناك حاجة لإرسالها عبر الإنترنت مع كل برنامج أبليت، يضاف إلى ذلك أنه يمكن كتابتها برمز الآلة لكل حاسب مستقل، وعليه يمكن تنفيذها أسرع مما ينفذ البايتكود المُفسّر (المترجم).

إن لغة الجافا يمكن أن تقدم الكثير لمجتمع المكتبات الرقمية، إلا إنها قد تنطوي على بعض أوجه القصور، وبعض عيوبها هي عيوب فنية تقليدية،

فهي لغة ضخمة، ومعقدة، وصعبة التعلم. واللغات المفَسرة عادة ما تنفذ ببطء، وجافا ليست مستثناة من ذلك، يضاف إلى ذلك أن قرارات التصميم التي تمنع تطبيقات جافا أبليت من نقل الفيروسات عبر الشبكة، وإصابة حاسبات المستفيدين، تعمل كذلك على إعاقة البرامج المشروعة. ومع ذلك فإن أكبر عيب لجافا ليس عيباً فنياً، بل إن مطورها، وهو شركة سن ميكروسيستم Sun microsystems، اتجهت نحو تطوير لغة قياسية، لكن لسوء الحظ فإن الشركات الأخرى، وخاصة مايكروسوفت، قد أوجدت بدائل غير متوافقة مع هذه اللغة.

تطورات حديثة في تصميم واجهات المستفيدين :

ينطوي تصميم واجهات المستفيدين على الطابع الجمالي، والطابع العلمي. وبشكل عام يعتمد نجاح واجهات المستفيدين على فراسة المصممين ومدى وعيهم بطبيعة سمات المستفيدين. ومن الملاحظ أن ثمة أفكاراً جديدة تُثري بشكل مستمر توجهات تصميم واجهات المستفيدين. ويلقى هذا الجزء من الكتاب على بعض هذه الأفكار الجديدة، وعلى بعض الموضوعات البحثية في هذا الجانب. كما سيتم في الفصل الثاني عشر عرض للاتجاهات البحثية في مجال ما وراء البيانات البنائية الذي يكتسب أهمية كبرى بالنسبة لواجهات المستفيدين؛ لأن الوضع الذي تشكّل فيه الكائنات الرقمية وما وراء البيانات البنائية التي ترتبط بذلك الوضع، يقدمان المادة الخام التي يعتمد عليها في تصميم واجهات المستفيدين.

النماذج التصورية CONCEPTUAL MODELS :

ظلت كثير من المجموعات البحثية تنقب عن النماذج التصورية التي

تساعد المستفيدين على الملاحة عبر المجموعات الهائلة من المواد المتاحة الآن مباشرة على الإنترنت. وإذا كان قليل من المعالم الجغرافية والخرائط والصور هو المتاح الآن على الإنترنت، فإن الروابط الفائقة التي هي قلب الويب توصل إلى مواقع غير متوقعة، مع أن المستفيد يمكن أن يضل طريقه في الوصول إلى تلك المواقع بسهولة. وإذا كان المستفيدون من المكتبات الرقمية قلما يتلقون برامج تدريبية رسمية، وإنما يعملون بمفردهم، فإن ذلك يؤكد على ضرورة تصميم واجهات تفاعل تعتمد على نماذج تصورية من أجل توجيه المستفيدين إلى الطرق الملائمة للوصول على المعلومات، هذا على الرغم من أن بعض المستفيدين وبشيء من الإصرار أصبحوا خبراء في كيفية الحصول على المعلومات من الإنترنت. وقد أشارت بعض الدراسات إلى أن أصحاب الخبرة يستطيعون إيجاد نسبة كبيرة من إحتياجاتهم من المكتبة مع وجود المعلومات على الشبكة في حين يضل قليلو الخبرة طريقهم في الوصول إلى المعلومات، بل ويجدون صعوبات كثيرة في تقييم المعلومات التي يحصلون عليها. وتقدم اللوحة رقم (٨-٣) وصفاً لمشروعين بحثيين استكشاف نماذج تصورية جديدة للمكتبات الرقمية.

اللوحة رقم (٨ - ٣)

نماذج تصورية جديدة : دي ليت **Dlite** وباد بلس بلس **Pad++**

لقد أجريت العديد من التجارب المثيرة حول صور واجهات المستفيدين بعيداً عن سطح المكتب، ولأن هذه التجارب تدخل في إطار المشروعات البحثية، فإن الواجهات التي تتم في سياقها لا تدخل في نظام الإنتاج، ومع ذلك ستظل أهمية هذه الجهود في تطوير مفاهيم يمكن الإفادة منها عند تصميم واجهات تفاعل أخرى، ولتوضيح عملية التصميم التي تعتمد على

التحليل المنهجي لاحتياجات المستفيدين وتوقعاتهم.

ودي ليت Dlite واجهة تفاعل تجريبية طورها ستيف كوزينز Steve Causins في جامعة ستانفورد جزءًا من مبادرة المكتبات الرقمية. وبحكم تطويرها واجهة تفاعل لنظام جامعة ستانفورد المسمى بأستانفوردإنفوبص standard Info Bus، تستخدم هذه الواجهة مفاهيم عملية البرمجة الموجهة بالهدف Object-oriented programming، وقد تم تنفيذ كل مكون من مكونات هذه الواجهة على أنه كائن كوربا CORBA (ويتضمن الفصل الثالث عشر وصفاً لكل من أنفوبص وكوربا (InfoBus, CORBA).

ويعتمد النموذج التصوري لدى (دي ليت) على تحليل الأنشطة التي يقوم بها المستفيد من المكتبة الرقمية، وقد تم تحديد المتطلبات التالية:

- تتكون المكتبة الرقمية من مجموعات متنوعة يجب الوصول إليها من أي مكان على الإنترنت.
- يمكن للنتائج التي وفرتها إحدى الخدمات أن تصبح مدخلات لخدمات أخرى.
- يجب أن تكون واجهة المستفيدين قابلة للتوسع، بحيث يمكن للمصادر الجديدة أن تتكامل بسهولة مع المصادر الموجودة من قبل.
- يمكن للمصادر أن تبقى لفترة زمنية طويلة.
- يجب أن يكون لدى المستفيدين القدرة على التعاون.

ويصف نموذج (ديليت) المكتبات الرقمية من حيث مكوناتها بأنها أربعة مكونات رئيسة هي: الوثائق، والاستفسارات، والمجموعات، والخدمات.

وهذه المكونات الأربعة يتم عرضها في شكل أيقونات يمكن معالجتها مباشرة من خلال نوافذ العرض المتاحة للمستفيد على الشاشة. حيث يمكن على سبيل المثال أن يؤدي طرح استفسار معين على خدمة البحث إلى إجراء عملية هي عملية البحث، ومن ثم الوصول إلى مجموعة من النتائج. ومما تجدر الإشارة إليه أن نموذج (ديليت) يسمح للمستفيدين بإنشاء واجهات تفاعل محددة المهام task-specific interface، عن طريق إظهار جميع عناصر الواجهة على الشاشة.

وثمة اعتباران شجعا على تبني التصميم الوظيفي لنموذج (ديليت) هما: سرعة الاستجابة، والقدرة على إضافة خدمات جديدة بأقل جهد ممكن. وبناءً على هذين الاعتبارين تم تقسيم التصميم الوظيفي إلى الحاسبات العميلة لواجهات المستفيدين user-interface client والحاسب الخادم لواجهة المستفيد user-interface server؛ بحيث تقوم الحاسبات العميلة بإجراء العمليات بين المكونات، والتي يمكن تنفيذ العديد منها في الوقت نفسه، أما الحاسب الخادم فيقوم بتوفير واجهات تفاعل للخدمات الخارجية، ويمكن لهذا الخادم أن يستمر في أداء عمله حتى عندما تكون الحاسبات العميلة خارج الخدمة.

أما القابلية للتوسع فيوفرها الحاسب الخادم؛ فبمجرد أن تضاف خدمة جديدة، تتم برمجة واجهة جديدة على الحاسب الخادم، أما الحاسبات العميلة فلا تحتاج إلى أية تعديلات. كما أن دعم مجموعة متنوعة من الحاسبات ونظم التشغيل المختلفة يقدم من خلال توفير برامج حاسبات عميلة مستقلة لكل نوع منها.

أما النموذج الآخر المسمى (باد بلص بلص pad ++)، فهو تصور لواجهة

تفاعل كان قد طرحه كين بيرلين Ken Perlin في جامعة نيويورك، ثم تبنى تطويره عدد كبير من الباحثين في كثير من الجامعات ومراكز البحوث. وتتمحور فكرة هذا التصور في عرض مجموعة كبيرة من المعلومات بأحجام أو مقاسات مختلفة. وهذا النموذج يتعامل مع صورة سطح المكتب أو يأخذها إلى ما وراء حدود شاشة الحاسب الآلي، ويتعامل معها كما لو كانت مجموعة من الوثائق المعروضة على جدار كبير. ويرى أن تفاعلات المستفيد تستند إلى الأفكار المألوفة لعملية التقريب والتبعيد والتحريك pamming and zooming؛ فالمستفيد يستطيع أن يقوم بتبعيد الصورة ليرى كل المجموعة ولكن بتفاصيل محدودة، أو يقوم بتقريبها جزئياً ليتمكن من رؤية جزء من المجموعة، أو بتقريبها ليرى أدق التفاصيل. وبما أن المستفيدين يتمتعون بذاكرة مكانية جيدة spatial memory، فقد ركز هذا النموذج على كل من الهيئة أو الشكل shape والمكان واعتبرهما مؤشرات تساعد المستفيدين على اكتشاف المعلومات وعلى استدعاء ما وجدوه بعد ذلك.

وعندما يقوم النموذج بعملية عدم تركيز زاوية الكاميرا Zoom out فإن التفاصيل لا تظهر، ولو حدث وظهرت فإن البرنامج سيضطر إلى إحضار أعداد هائلة من التفاصيل الدقيقة جداً، ولا شك أن لكل ميزة ما يعيبها. وعند مقياس معين فإن هذه المزايا تتداخل مع مزايا أخرى، أو أنها لا تظهر على الإطلاق، وفي عملية تركيز بؤرة الكاميرا الدلالية semantic zomming تغير الكائنات مظهرها عندما يتغير حجمها، وهو أسلوب مألوف في علم الخرائط، حيث يمكن أن تظهر المدن في خريطة كبيرة، ولا تظهر بالطبع

المباني الموجودة في هذه المدن.

ومما تجدر الإشارة إليه أن نموذج (باد بلص بلص) لم يكن القصد منه أن يكون واجهة تفاعل شاملة لكل التطبيقات، حيث يمكن لهذا النموذج أن يكون بمثابة واجهة تفاعل كاملة في بعض الأغراض مثل استكشاف مجموعات كبيرة متدرجة بشكل هرمي في المكتبات الرقمية، لكنه في بعض الأغراض الأخرى، يعتبر أداة تخيل أو رؤية visualization component، تضاف إلى جانب المفاهيم الأخرى لواجهات المستفيدين التي يتم التعامل معها بداخل نظام النوافذ التقليدي.

إن نموذج (باد بلص بلص) يعد مثالاً هاماً لبحوث واجهات المستفيدين التي تركز على دراسة العلاقة بين أداء النظام ومدى إمكانية استخدامه. ولقد تم تطوير إصدارات عديدة من نموذج (باد بلص بلص) يعمل بعضها بشكل مستقل، في حين يعمل الآخر منها كجزء من متصفحات الويب. كما يلاحظ أن كل إصدارة من هذه الإصدارات تحتوي على إمكانية مراقبة الأداء الداخلي. فعمليات العرض عادة ما تكون مؤقتة لأن معدل تنشيط الإطار يظل ثابتاً أثناء عمليات تبعيد زاوية الكاميرا وتقريبها وتحريكها، وعندما تبدأ الواجهة في التباطؤ، فإن الخصائص متوسطة الحجم تعرض بشكل متوسط، كما أن التفاصيل تضاف عندما يعمل النظام ببطء.

تصميم الواجهة :

بالرغم من البعد الجمالي الذي يغلب على توجهات تصميم واجهات المستفيدين، فإن ثمة عدداً من المبادئ العامة الأخرى التي ينبغي مراعاتها في هذا الصدد، والتي خلصت إليها جهود البحث العلمي المتصلة بهذا الجانب في

الآونة الأخيرة. وهى مبادئ لها أهميتها الخاصة بالنسبة للمستفيدين، ومن هذه المبادئ الثبات والتناغم في عناصر المظهر وفي عمليات التحكم فيه وضبطه بالإضافة إلى الوظائف التي تؤدى.

يضاف إلى ذلك أن المستفيدين دائماً يحتاجون إلى التغذية المرتجعة، فهم يحتاجون إلى فهم ما يقوم به نظام الحاسب الآلي من أعمال ولماذا يرون نتائج محددة. كما يتطلعون إلى إتاحة الإمكانية لهم في تعطيل بعض الأعمال التي يقوم بها الحاسب أو تغير مسارها، وأن تتوافر لهم إمكانات التعامل مع الأخطاء ومعالجتها على نحو مبسط وميسر. وفي الوقت الذي يكتفى فيه بإعطاء المستفيدين المهرة لمحات سريعة لكيفية التعامل مع النظام، يجب أن يتم توفير خيارات سهلة وواضحة جداً للمستفيدين المبتدئين. وفوق ذلك كله لابد أن يشعر المستفيد أن العمل كله يخضع لسيطرته.

ودائماً ما تسبب عملية التحكم، وخاصة في التصميم الجرافيكي وفي مظهر المواد توتراً بين كل من مصممي المكتبات الرقمية والمستفيدين منها. فكثير من المصممين يطمح في أن يتمكن المستفيدون من مشاهدة المواد على صورتها التي صممت بها بالضبط، كما أنهم يريدون أن يتحكموا في درجة الدقة الجرافيكية والطبوجرافية وحجم النوافذ وأماكن وضع المعلومات داخل النافذة وغيرها من الجوانب الأخرى التي تتصل بأساسيات التصميم الجيد، ولكن لسوء حظ هؤلاء المصممين، فالمتصفحات عادة ما تكون أدوات عامة. والمصمم لا يعرف بالتحديد المتصفح الذي يستخدمه المستفيد، ولا نوع الحاسب الذي يستخدمه، ولا مدى سرعة الشبكة، وما إذا كانت نافذة العرض صغيرة أم كبيرة. كما أن المستفيدين قد يقومون بإعادة تشكيل حاسباتهم أو

تهيئتها، فقد يفضلون استعمال خطوط بأحجام كبيرة، أو نوافذ صغيرة، كما أنهم قد يرغبون في إغلاق شاشات عرض الصور بهدف التخفيف من حمل الشبكة وإسراعها. بناءً على ذلك فإن التصميم الجيد يجب أن يكون مناسباً للعمل في بيئات حاسوبية مختلفة، وأفضل المصممين هم من يتمتعون بموهبة تصميم واجهات تفاعل مناسبة من ناحية، وملائمة للعمل مع مختلف أنواع الحاسبات من ناحية أخرى، ومع ذلك فهناك بعض المصممين لا يزالون يجدون صعوبة في التحول من عالم الوسائط التقليدية - الذي يتحكمون في كل شيء فيه - إلى عالم المكتبات الرقمية والويب. وثمة خطأ شائع، وهو أن الإفراط في الإيضاح، قد يؤدي إلى عدم استخدام إحدى الواجهات إلا في ظل الشبكات السريعة ومع الحاسبات عالية الأداء.

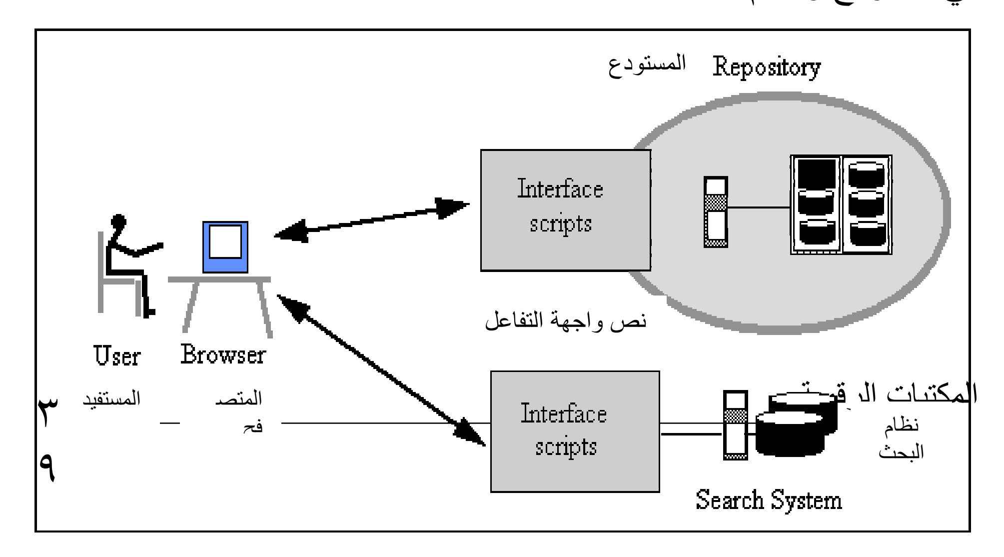
ومن التعقيدات غير المبررة، أن كلاً من نيتسكيب ومايكروسوفت - وهما من كبار موردي المتصفحات - تطرح منتجات اختارت لها أن تكون مختلفة عن المنتجات التي تطرحها الشركات الأخرى، وإذا أراد المصمم أن يستخدم بعض المزايا الخاصة فلا بد من إحاطة المستفيد علماً بأن هذا التطبيق لا يعمل مع جميع المتصفحات. وعلى ذلك فإن الواجهة التي قد تبدو سلسة في العمل مع أحد المتصفحات، قد تبدو معقدة مع متصفح آخر، ذلك أن روح المنافسة بين المنتجين قد وضعت في أولوياتها أشياء أخرى غير رضا المستفيدين.

التصميم الوظيفي :

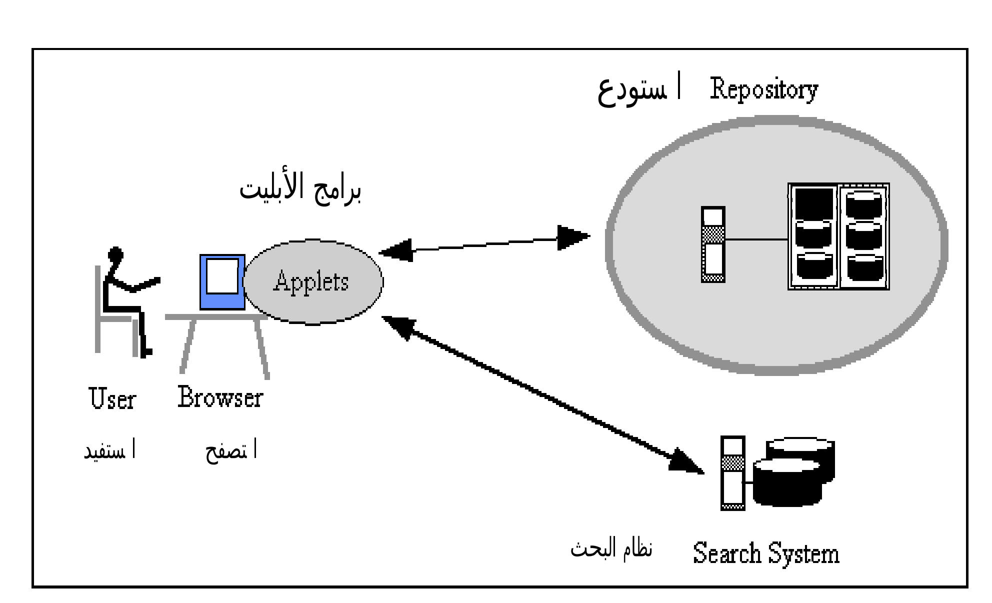
المكتبات الرقمية هي أنظمة موزعة، تتكون من العديد من الحاسبات التي يعمل بعضها مع بعضٍ كفريق عمل. ولاشك أن البحث في [جزئية]

التصميم الوظيفي يمد المصممين بخيارات عن وظيفة كل حاسب من هذه الحاسبات المختلفة، وعن العلاقة بين هذه الحاسبات كذلك. وفيما بين الحاسب المستودع repository (حيث يتم تخزين المجموعات) والمستفيد النهائي (المستخدم لمتصفح الويب) تتواجد مجموعة متنوعة من البرامج التي تعرف غالباً بالمكونات الوسيطة middleware. وبرمجيات المكونات الوسيطة هذه تعمل كحلقة وصل بين المستفيد من ناحية والمستودع من ناحية أخرى. حيث تقوم بترجمة التعليمات التي يرسلها المستفيد، ثم إرسالها إلى المستودعات، ثم تتلقى المعلومات من المستودعات، وتعمل على تنظيمها، ثم تعيد إرسالها إلى حاسبات المستفيدين. وغالباً ما تقدم هذه المكونات الوسيطة خدمات داعمة مثل توفير ضمانات الثقة، وتصحيح الأخطاء، ولكن أهم مهمة تقوم بها هي التوفيق بين الخدمات التي يطلبها المستفيد مع تلك التي يقدمها المستودع.

ويصور الشكلان (٨-٢)، (٨-٣) نموذجين شائعين، يظهر في كل منهما أن ثمة مستفيداً يجلس أمام حاسب شخصي مستخدماً أحد متصفحات الويب القياسية كواجهة تعامل أمامية للحصول على خدمات المكتبة الرقمية المتمثلة في مستودع ونظام بحث.



الشكل رقم (٨-٢) واجهة تفاعل تعمل بأسلوب الواجهة



الشكل رقم (٨-٣) واجهة تفاعل تعمل بأسلوب الرمز المتنقل (برامج الأبليت)

وفي الشكل رقم (٨-٢) يتم تنفيذ المكون الوسيط بمثابة أحد البرامج الحاسوبية المستقلة (عادة ما تكون نصوص بوابة العبور المشتركة CGI scripts) التي يتم تشغيلها على حاسبات خادمة موجودة في مكان ما على الشبكة، حيث يقوم المتصفح بإرسال الرسائل إلى واجهة المستفيدين عبر أحد البروتوكولات القياسية (ربما يكون بروتوكول نقل الملفات الفائقة HTTP)، ومن ثم يمكن أن تشغل النصوص على الخدمة البعيدة، أو يمكنها أن تتصل بالخدمة عبر وسائل معينة من خلال أي بروتوكول آخر مناسب. ويمكن في

هذه الحالـة أن يستخدم بـروتوكول المواصفة القياسية رقم Z39.50 للربط بين نـصوص الـواجهة ونظـام الـبحث. وبالـرغم مـن أن الاخـتلافات الفنـية لهـا أهميـتها بالنـسبة لمـرونة الـنظام وأدائـه؛ إلا أنهـا تعـد مـن الناحـية الوظيفـية مـتوازية. ويعيب هذا التشكيل أن عملية من العمليات التي يقوم بها المستفيد يجب أن ترسل عبر الشبكة لمعالجتها، ومن ثم يتعين على المستفيد أن ينتظر إلى حين استجابة النظام.

أمـا الـشكل رقـم (٨-٣) فيعـرض تـصميماً آخـر أكثـر تطـوراً، حيث يتم فيه تنفيذ الرمـز المتـنقل على حاسب المستفيد عن طريق المتصفح الذي يستخدمه، مـع ملاحظـة أن هـذا الرمـز قـد يكـون أحـد البـرامج التطبيقـية الـصغيرة لجافـا [جافـا أبليت]، وبمجرد استخدام هذا الرمز يتم اختزانه على الخادم البعيد، ويتم تحميله علـى حاسـب المستفيد تلقائياً عند الطلـب، ولـهذا التـشكيل عدة مـزايا، فحيث إن رمـز واجهـة المستفيد يتم تشغيله على حاسب المستفيد، فإن استجابته تكـون سـريعة، كمـا أن البـروتوكول الـذي يربط بين واجهة المستفيد والخدمات يمكـن أن يكـون بـروتوكولاً قياسياً أو أي بـروتوكول آخـر يـتم إعداده خصيصاً لـهذا التطبيق.

وثمــة تـصور جـيد ظهـر فـي الآونــة الأخيـرة يعـرف بملـف العـرض presentation profile يقـوم مديـرو المكتبات الـرقمية بموجبه بـربط التوجيـهات بالمعلـومات المختـزنة، علـى أن تقوم هذه التوجيهات باقتراح الشكل الذي يجب أن يعـرض بـه الكـائن الرقمـي object للمستفيد، حـيث يمكـن لملـف العـرض - علـى سـبيل المـثال - أن يقتـرح للمُستفيد طـريقتين: تقدم إحـداهما ملفاً صـغير

الحجم للكائن، في حين تقدم الثانية التفاصيل الكاملة. وعادة ما ينصح أن تتبع واجهة المستفيد أسلوب ملف العرض في عرض البيانات وإن كان لها الخيار في اتباع أية طريقة أخرى مناسبة.

أنظمة الحاسبات والشبكات :

لأنظمة الحاسبات والشبكات أثرها الكبير على مدى الإفادة أو الاستخدام، فمنشئ المكتبة الرقمية قد يضع افتراضات قليلة حول التجهيزات التي تتوافر للمستفيد، بعيداً عن المعرفة الأساسية بأن لكل مستفيد حاسبًا شخصيًا متصلاً بالإنترنت. وبالرغم من ذلك، فهذه العبارة البسيطة تغطي مجموعة متعددة من الحالات، فهناك حاسبات شخصية بها إمكانات أكبر من بعض الحاسبات الأخرى، كما أن جودة أجهزة العرض تختلف اختلافاً كبيراً من حاسب آلي لآخر، يضاف إلى ذلك أن يكون لبعض المستفيدين حاسباتهم الخاصة التي يتحكمون في تهيئتها أو تشكيلها دون قيود، في حين يتقاسم آخرون العمل على حاسب واحد أو يستخدمون أكثر من حاسب واحد. والقائمون على كل مكتبة رقمية يجب أن يضعوا في اعتبارهم هذا التفاوت في إمكانات الحاسبات التي يستخدمها المستفيدون مع التفاوت في نظم التشغيل وفي بيئات العمل التي تعمل فيها هذه الحاسبات المختلفة. ومن صور بيئات العمل المختلفة تلك الإصدارات المتتالية لنظم النوافذ لميكروسوفت، ولنظام الماكنتوش، ومختلف أشكال نظام اليونكس. مع ملاحظة أن أبعاد الاختلافات بين نظم التشغيل هذه يمكن التقليل من حدتها عن طريق استخدام أحد متصفحات الويب لواجهة المستفيد، ومع ذلك فليس من السهولة معالجة الفوارق في معدلات الأداء.

وإذا كان مدى استخدام أحد نظم الحاسبات يعتمد على مدى سرعته في الاستجابة للتعليمات، فقد لا يكون لمصمم المكتبة الرقمية معرفة كافية، أو قل لديه معرفة محدودة، بمدى جودة توصيلات الشبكة التي تربط بين كل من المستفيد والمكتبة، فقد تتفاوت هذه التوصيلات تفاوتاً كبيراً من حيث سرعتها، ما بين السرعة الفائقة والبطء الممل الذي يدعو للضجر، مع ملاحظة أن هذه التوصيلات السريعة قد لا تسلم من الأعطال، كما يلاحظ كذلك أن واجهة المستفيد التي تدعو إلى البهجة عند استعمالها مع شبكة محلية ذات سرعة تصل إلى أكثر من عشرة ملايين بت في الثانية، قد لا يمكن استعمالها مع خط اتصال غير مألوف لا تزيد سرعته القصوى على ٢٨٠٠٠ بت في الثانية، وبناءً على ذلك، يجب على القائمين على المكتبات الرقمية أن يوازنوا بين الاستخدام الفعال للخدمات المتقدمة، التي تتطلب تجهيزات سريعة وبرمجيات حديثة جدًا والخدمات الجيدة التي تقدم للمستفيدين الذي لم يسبق أن قدمت لهم من قبل.

تحسين درجة استجابة المتصفحات :

تحاول متصفحات الويب القيام ببعض الخدع التي من شأنها رفع درجة الاستجابة التي يلحظها المستفيد، ومن إحدى هذه الخدع ما يعرف بالإخفاء الداخلي أو بالذاكرة المخبأة الداخلية internal caching، وتستند هذه الطريقة على افتراض مفاده أن المعلومات التي يتصفحها أحد المستفيد في بعض الأوقات قد يحتاج للرجوع إليها بعد فترة قليلة. (على سبيل المثال عندما يقوم المستفيد بالضغط على زر "العودة للخلف" back أثناء عمله على أحد المتصفحات)، كما أن هناك بعض الرسوم التصويرية (مثل العلامات والأزرار) قد يتكرر استخدامها باستمرار، بناء على هذا الافتراض تقوم

متصفحات الويب بالاحتفاظ المؤقت للملفات المستخدمة مؤخراً عن طريق تخزينها مؤقتاً على القرص الخاص بالحاسب الشخصي للمستفيد، وهذه الملفات قد تكون عبارة عن صفحات مهيأة بلغة تهيئة النصوص الفائقة HTML، أو تكون صوراً، أو ملفات الرمز المتنقل (التي قد تكون مكتوبة بنصوص جافا أو جافا سكريبت)، وعندما يقوم المستفيد بطلب أحد هذه الملفات، فإن المتصفح يقوم أولاً بالبحث عنه للتأكد من إمكانية إعادة قراءته من الذاكرة المخبأة الداخلية بدلاً من محاولة استرجاعه من خادم بعيد.

ومن الأساليب الأخرى لتحسين درجة استجابة المتصفحات القيام بإجراء عدة عمليات في الوقت ذاته على نحو متوازٍ، حيث تقوم المتصفحات بعرض الجزء الأول من أحد الملفات الطويلة قبل وصول الملف كاملاً، كما يتم عرض الصور كمخططات مختصرة أولاً، ثم تتم بعد ذلك تعبئة الفراغات بعد وصول البيانات، كما يمكن طلب عدة أشكال مختلفة من البيانات في آن واحد بصورة متوازية، وإجمالاً يمكن القول إن هذه الأساليب تعمل على التخفيف من حدة بطء الأداء وضعفه اللذين عادة ما يصيبان الإنترنت.

الذاكرة المخبأة وذاكرة المواقع المطابقة mirroring and caching :

تعد عملية استنساخ البيانات من الأساليب شائعة الاستخدام على نطاق واسع، وهي تهدف إلى تحسين أداء النظم على الإنترنت.

فإذا أراد أناس في مختلف أنحاء العالم استخدام إحدى المكتبات الرقمية، فمن الممكن وضع نسخ مكررة من المجموعة الكاملة لهذه المكتبة في عدة مواقع - ربما تكون اثنتان منها في أوربا، واثنتان أخريان في الولايات

المتحدة الأمريكية، وواحدة في أستراليا، وأخرى في اليابان ـ وهذه العملية تعرف بعمل المواقع المطابقة [البديلة] mirroring وعادة ما يختار المستفيدون بشكل عام أقرب موقع مطابق يقدم لهم أفضل أداء، وتسمح عملية المواقع المطابقة هذه للمستفيد بالتوجه إلى موقع آخر في حال تعطل الموقع الذي توجه إليه في البداية، أو إذا ما واجه مشكلة ما في الشبكة.

تُعرف عملية استنساخ معلومات محددة (وليكن المكتبة الرقمية برمتها) بعملية التخبئة أو الإخفاء caching. والذاكرة المخبأة cache هي عبارة عن مستودع يحتفظ بالمعلومات التي استخدمت في الآونة الأخيرة، تجنباً لتأخر الاستجابة عند استخدام تلك المعلومات مرة أخرى. ومستودعات الذاكرة المخبأة تتواجد بداخل التكوين المادي للحاسب وتستخدم بهدف مساعدة معالج الحاسب في العمل بشكل سريع، كما أنها تتواجد في أدوات التحكم أو المراقبات controllers، والتي تقوم بقراءة البيانات من أقراص الحاسب، أما بالنسبة لنظام اسم المجال أو النطاق domain-name فتستخدم الذاكرة المخبأة بهدف تحسين درجة السرعة والثقة في تحويل أسماء المجالات إلى عناوين الإنترنت.

وللمكتبات الرقمية ذاكرات مخبأة في العديد من الأماكن؛ حيث يمكن لإحدى المؤسسات مثلاً أن تحتفظ بذاكرات مخبأة داخلية للوثائق الخاصة بها والتي تمت قراءتها في الآونة الأخيرة، كما ستقوم المكتبات الرقمية التي تختزن المجموعات الكبيرة من البيانات على أجهزة تخزين بطيئة، ولكنها رخيصة الثمن، بالاحتفاظ بالذاكرات المخبأة على أجهزة تخزين أسرع.

وتجدر الإشارة إلى أن جميع وسائل استنساخ البيانات تعاني من التفاوت

بين الإصدارات المختلفة، واحتمال أن يتلقى المستفيدون مواد متفاوتة، ولذلك فلا بد من اتخاذ إجراءات تحديث تعمل على استبعاد المعلومات التي تم استنساخها بعد وقت محدد، أو تعمل على فحص المصدر للتعرف إلى مدى حدوث أية تغيرات.

ومما يذكر أن مواضع الذاكرة المخبأة هذه لا تسلم من التعرض للاختراق وأعمال القرصنة، وإذا كان أمن النظم الآلية يقاس بأمن أضعف حلقاتها، فإن الذاكرة المخبأة تعد أضعف تلك الحلقات.

مدى الثقة وواجهات المستفيدين :

قليلة هي النظم الآلية التي يمكن التعويل عليها بشكل كامل، والمكتبات الرقمية تعتمد على عديد من النظم الفرعية الموزعة عبر شبكة الإنترنت، وبالنظر إلى عدد المكونات المنفصلة بعضها عن بعض، يكون من الأشياء المعتبرة أن يعمل أي نظام عالمي.

وتحرص النظم الآلية التي أحكم تصميمها على إحاطة المستفيد علماً بما يشير إلى تقدم سير العمل أثناء أدائها لمهمتها، ومن أبسط وسائل هذه الإحاطة ظهور صورة متحركة أثناء انتظار الاستجابة، وهو ما يعطي المستفيد انطباعاً بأن ثمة شيئاً ما يتم، أما أكثر هذه الوسائل تقدماً فيتمثل في ظهور مؤشر يبين إنجاز المهمة وتقدير الوقت المتبقي لإكمالها، وفي جميع هذه الحالات، يجب أن يمنح المستفيد إمكانية إلغاء أي عملية يراها ستستغرق وقتاً طويلاً.

وهناك مصطلح "التناقص اللبق" graceful degradation الذي يصف

الوسـائل التـي تبين أن المهمـة المطلـوب إنجازهـا ستستغرق وقتاً طـويلاً، مـع محاولـة تقديم وسـائل إرضـاء جزئية للمستفيد، ومن أبسط الوسائل المستخدمة فـي هذا الـصدد تمكين مستخدمي المتصفح من إغلاق الصور التي تظهر على صفحات الـويب، والاكتفاء بـرؤية النـصوص فقط. كمـا تقوم كثير من أساليب بث الـصور أو توصيلها بعرض صـورة غير دقيقة بمجرد وصول جزء من البيانات في أول الأمر، ثم التدرج في إكمال دقة وضوح الصورة رويداً رويداً إلى حين اكتمالها بعد ذلك.

واجهات المستفيدين لمعلومات الوسائط المتعددة :

لقد أصبحت المواد المتاحة في المكتبات الرقمية متنوعة بشكل كبير في نوعياتها وأشكالها. وربمـا لا تكفـي واجهـة المستفيدين المصممة للتعامل مع المجموعات التـي تحتوي علـى مواد نصية للتعامل مع مجموعات الموسيقى والخـرائط وبـرمجيات الحاسـب، والـصور، والبـيانات الإحـصائية وألعـاب الفيديو، ومع ذلك، فهناك بعض المبادئ العامـة التـي يمكن تطبيقها في هذا الصدد:

- التلخيص أو العرض المختصر summarization :

من الملائم للمستفيد عندما تتاح لـه عدة مواد ليختار منها ما يناسبه أن تمـثل كـل مـادة مـن هـذه المـواد بـبعض المعلـومات المختـصرة التـي توضـح مضمونها؛ فالكتاب عادة ما يتم عرضه مختصراً عن طريق ذكر اسم مؤلفه وعنوانه (هكذا مثلاً Arms, Digital library)، أمـا الـصورة فيمكن الاكتفاء بعرض صـورة منمنمة منها، لكن علـى النقيض؛ إذ يصعب علـى المكتبات

الرقمية تلخيص المعلومات ذات الطبيعة الأكثر تعقيداً مثل برامج الحاسب الآلي والأفلام.

– الفرز والتقسيم إلى فئات التصنيف sorting and categorization :

تنظم المكتبات المعلومات عن طريق تصنيفها أو ترتيبها وفق تسلسل معين، غير أن هناك أشكالاً أخرى عديدة من المعلومات (مثل النغمات الموسيقية) ليس لها طريقة مبسطة ومفهومة لتنظيمها. وإذا كان من السهل استخدام أدلة إحصاء الطيور من حيث تقسيمها إلى فئات معروفة، مثل: البط، والصقور، والعصافير، فإن من الصعب استخدام أدلة الزهور البرية لأن ليس لمثل هذه الزهور من تقسيمات متوازية.

- العرض : تحتاج أشكال كثيرة من المعلومات إلى تجهيزات خاصة لعرضها للمستفيد بشكل جيد. فالصور المتحركة الرقمية عالية الجودة تحتاج في عرضها إلى أجهزة عرض خاصة، وحاسبات بإمكانات قوية، وشبكة اتصال فائقة السرعة، وبرمجيات مناسبة. ويجب على المكتبات الرقمية أن تعمل على الإفادة الكاملة من مثل تلك الإمكانات، إن وجدت، وأما بالنسبة لعرض تلك المعلومات للمستفيدين الذين لا يتوافر لهم إلا القليل من مثل تلك التجهيزات، فلا بد من العمل على توفير وسائل عرض بديلة.

وليست هناك طريقة وحيدة لتصميم مكتبات رقمية تتعامل مع مختلف أشكال المعلومات. وعليه يجب أن تعد واجهات المستفيدين التي تتعامل مع مختلف أشكال المواد، وربما كذلك مع مختلف فئات المستفيدين. وتصف اللوحة رقم (٨-٤)، مثالاً جيداً لإحدى واجهات المستفيدين: وهي مكتبة

إنفورميديا الرقمية للمقاطع المرقمنة من الصور المتحركة The informedia digital library of digitized segment of video.

اللوحة رقم (٨-٤)

إنفورميديا **INFORMDIA**

إنفورميديا هو أحد البرامج الموجودة في جامعة كارنيجي ميلون قام بإدارته هوارد واكتلار Howard Wacttlar، وهو برنامج بحثي عن المكتبات الرقمية الخاصة بالمواد المتحركة أو مواد الفيديو، وخاصة نشرات الأخبار والبرامج الوثائقية، وقد جزئت البرامج تلقائياً إلى مقاطع قصيرة، مثل المواد الفردية في إحدى نشرات الأخبار. وينصب تركيز هذا البرنامج على تطوير وسائل آلية لاستخلاص المعلومات من الكائنات الحركية Video object بهدف الحد من التدخل البشري في ذلك. وقد استخدم لهذا الغرض أكثر من ألف ساعة من مواد الفيديو الرقمية (التي تم الحصول عليها من شبكة سي إن إن الإخبارية، والجامعة البريطانية المفتوحة، ومحطة التلفزة العامة في بتسبرج والمعروفة بـ WQED وغيرها من المصادر الأخرى).

وتتمثل إحدى القضايا المتصلة بمواد الفيديو أو المواد السمعية في كيفية إمداد المستفيدين بعرض سريع لإحدى هذه المواد، فإذا كان القارئ لأحد الكتب يريد أن يلقي نظرة على صفحة المحتويات، أو يتصفح صفحات الكتاب ليستكشف بنفسه العناوين الرئيسية لفصوله، فإن مثل تلك الإمكانات لا تتوافر في حال مواد الفيديو. وقد عمل مشروع إنفورميديا على تطوير طريقة لتصفح صور الفيديو Video skimming، وتتمثل هذه الطريقة في توفير وسائل آلية لاستخلاص الكلمات والصور الهامة، وفي تركيبة واحدة تقوم الكلمات

والصور التي تم اختيارها بتقديم مستخلص فيديوي video abstract يوضح محتوى المقطع المصور كاملاً. ومن اللافت للنظر أن آفاق البحث في برنامج إنفورميديا حول عملية استخلاص الأفكار من صور الفيديو قد وصلت إلى أبعد مما كان متوقعاً لها عند بداية المشروع.

وتستخدم واجهة المستفيدين في مشروع إنفورميديا النموذج التصوري عند بحث كشاف المجموعة، والحصول على عدد من الإصابات أو النتائج، ثم التصفح لهذه المواد التي يتم الحصول عليها. أما بالنسبة للاستفسارات الموجهة للنظام فيتم إدخالها عن طريق كتابتها أو من خلال وسائل برنامج التعرف إلى بصمات الصوت .speech recognition وبعد عملية البحث تقوم الواجهة بعرض النتائج في ترتيب طبقي، بحيث يتم تمثيل كل مقطوعة فيديو بصورة معينة، وعندما يتحرك المؤشر فوق هذه الصورة، يتم إتاحة ملخص نصي لمحتواها، مع ملاحظة أن الصور التي تمثل تلك المقاطع يتم اختيارها تلقائياً بناءً على مدى ارتباطها بموضوع الاستفسار الذي قُدِّمَ، أما الملخص فيتم إعداده عن طريق عملية معالجة اللغة الطبيعية Natural language processing، ويمكن للمستفيد بعد ذلك ومن خلال الضغط على الصورة أن يرى المقطع كاملاً. وتجدر الإشارة إلى أن واجهة المستفيدين لهذا البرنامج تتكون من أدوات عرض صور الفيديو المصحوبة بوسائل التحكم (مثل "ابدأ"، و"توقف") كما هو الحال في جهاز الفيديو المنزلي. وتتمتع واجهة المستفيدين في هذا النظام بميزة سهولة استخدامها لدرجة أن من لم تتح لهم فرص تدريب سابقة على استخدام هذا النظام لا يجدون صعوبة في

استخدامها، كما أن هذه الواجهة يتوافر فيها ما يساعد على رفع معدلات التغذية المرتدة للمستفيد في كل خطوة يقوم بها، وأخيراً تجدر الإشارة إلى أن جميع الاختبارات التجريبية على هذا النظام أجريت على طلاب المرحلة الثانوية ببتسبرج.

واجهات المستفيدين وفاعلية المكتبات الرقمية :

لقد ارتقت معدلات جودة تصميم واجهات المستفيدين خلال السنوات القليلة الماضية ارتقاءً كبيراً، ويفترض الآن أن يبدأ المستفيدون الجدد عملاً مثمراً دون حاجة لأية برامج تدريبية، بل الأهم من ذلك، أنه يتوافر الآن على الإنترنت نماذج كثيرة لواجهات جيدة، يمكن للآخرين استخدامها أو الإفادة منها سواءً من أجل الاقتداء بها أو التعامل معها كنماذج للتصميم الجيد. يضاف إلى ذلك أن مواصفات التصميم الجرافيكي أو التصوري تشهد هي الأخرى تحسناً سنوياً. تجدر الإشارة إلى أن مواد مجلة المكتبات الرقمية DL magazine أعيد تصميمها ثلاث مرات خلال السنوات الماضية، وبالرغم من أن كل تصميم كان يبدو لنا رائعاً في حينه، إلا أنه يتبين لنا بعد سنة واحدة أنه بحاجة إلى عملية تجميل (Face lift).

إن الدعم الجيد للمستفيدين أكثر من مجرد عملية زخرفة تجميلية، كما أن التصميم الجيد، والأداء الوظيفي المناسب، والنظم سريعة الاستجابة هي التي تعد معيار القياس في فاعلية المكتبات الرقمية، وعندما تكون هناك صعوبة في استخدام آلي معين، فإن المستفيدين قد يفشلون في الوصول إلى النتائج الهامة، بل قد يخطئون في تفسير ما يصلون إليه، وربما ينصرفون

عن البحث بعد إصابتهم بالسأم، وأخيراً يمكن القول بأنه بقدر جودة واجهة المستفيدين تكون جودة المكتبة الرقمية.

الفصل التاسع
النص

للمواد النصية مكانة خاصة في جميع المكتبات، بما في ذلك المكتبات الرقمية. وعلى الرغم من أن الصورة قد تغني في بعض الأحيان عن ألف كلمة، فإن أفضل طريقة لنقل الأفكار المعقدة في أغلب الأحيان تكون عن طريق الكلمات، وإن ثراء المفاهيم والتفاصيل ودقة الأفكار التي يمكن التعبير عنها في النص لهي من الأمور التي تلفت الانتباه.

ومصادر الوثائق النصية في المكتبات الرقمية متعددة، فبعضها يتم إنشاؤه لأغراض الاستخدام المباشر online use، وبعضها الآخر يتم عن طريق تحويل المواد المطبوعة أو غيرها إلى مواد رقمية، كما أن بعضها قد يكون عبارة عن تسجيلات رقمية مستقاة من أفلام أو برامج تلفزيونية. وللسجلات النصية Textual records وظيفتها الخاصة في هذا السياق حيث تعد بمثابة ما وراء البيانات metadata التي تصف المواد الأخرى.

ويناقش هذا الفصل كيفية تمثيل الوثائق النصية لأغراض التخزين في الحاسبات، وكيف يتم إعادة تحويلها بعد ذلك عند طباعتها أو عرضها للمستفيدين. أما فيما يتصل

بتسجيلات ما وراء البيانات وأساليب بحث الوثائق النصية فقد أرجأنا الحديث عنه لفصول لاحقة.

التهيئة أو الترميز mark up ووصف الصفحات، ونماذج عرض الصفحات style sheet:

يجب الإشارة في البداية إلى أن طرق اختزان المواد النصية عليها أن تبين جانبين مختلفين من جوانب الوثيقة التي يتم تمثيلها، هذان الجانبان هما: هيكل أو بنية structure هذه الوثيقة، وشكلها أو مظهرها appearance.

ويصف الهيكل تقسيم النص إلى عناصر، مثل الحروف والكلمات والفقرات والعناوين الرئيسة، كما يحدد أجزاء الوثيقة التي يجب التركيز عليها، ويحدد كذلك الأجزاء التي يجب أن توضع في جداول أو حواشٍ، وأي شيء آخر يربط جزءاً بغيره من الأجزاء. وغالباً ما يتم ترميز بنية النص الذي يتم اختزانه في الحاسب عن طريق مواصفات الترميز أو التهيئة Markup Specification، وقد لاقت اللغة المعيارية الموحدة لتهيئة النصوص (SGML) في السنوات الأخيرة قبولاً واسعاً كنظام موحد لتهيئة بنية النصوص.

أما المظهر فهو وصف الحالة التي تبدو عليها الوثيقة عندما تُعرض على الشاشة أو عندما تطبع على ورق بعد ذلك. وعليه فإن المظهر أقرب ما يكون إلى عملية اختيار الشكل أو الصيغة format ، أي اختيار شكل حروف الطباعة وأحجامها، واختيار هوامش الصفحات والمسافات بين السطور، وكيفية عرض العناوين الرئيسة، ومواضع الأشكال والرسومات، وأساليب عرض المعادلات الرياضية وغيرها من الرموز الخاصة. مع ملاحظة أن القرارات المتعلقة بالمظهر في حالة الكتب المطبوعة قد تمتد لتشمل قرارات اختيار الورق ونوع التجليد. ويتم استخدام "لغات وصف الصفحات page-description languages" لأغراض اختزان الوثائق ونقلها أو معالجتها بطريقة تصف مظهرها وصفاً دقيقاً.

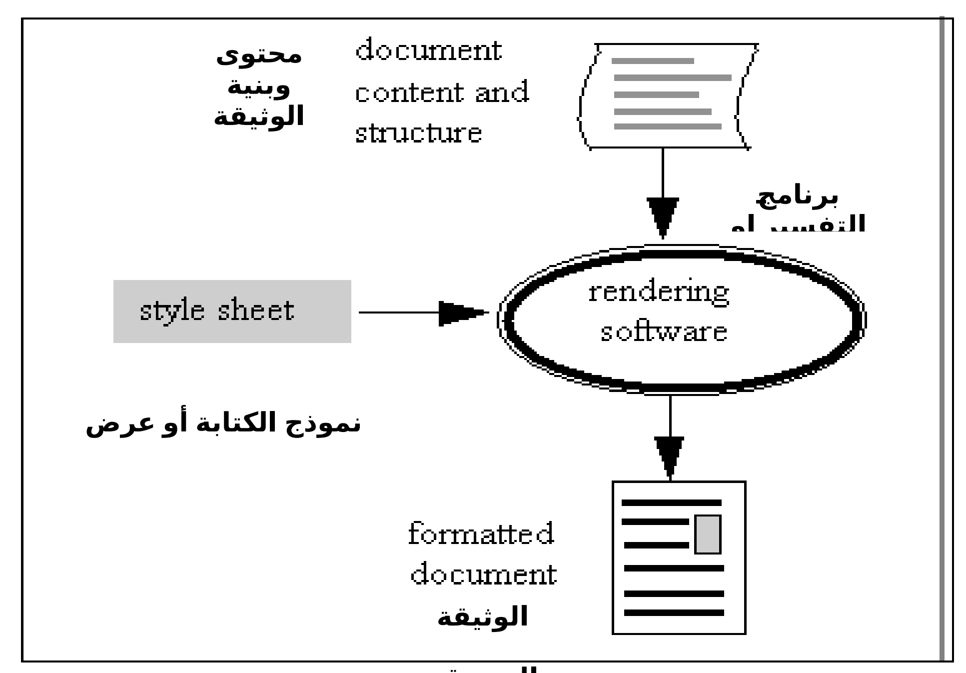
ويستعرض هذا الفصل ثلاثة أساليب مختلفة بعضها عن بعض إلى حد ما من أساليب وصف الصفحات، وهي: تي إكس TeX، وبوست سكريبت Post Script، وصيغة الوثيقة المحمولة PDF.

وتجدر الإشارة إلى أن الهيكل والمظهر مرتبطان بتصميم الوثيقة، فأثناء عملية النشر التقليدي يقوم المصمم

بوضع مجموعة من مواصفات التصميم التي تحدد الكيفية التي يجب أن يظهر بها كل عنصر من العناصر البنائية للوثيقة، مع قواعد شاملة لكل موقف من المواقف التي يمكن أن تظهر أثناء ذلك، وهذه المواصفات تمكن الجامع أو الطابع من إنشاء وثيقة منسقة تنسيقاً جيداً من مخطوطة قام بتهيئتها أحد المحررين أو المصممين أو كليهما.

ويوضح الشكل رقم (9-1) الإجراء الذي يلجأ إليه كثير من ناشري الدوريات عند إصدارهم للدوريات الإلكترونية. حيث يقومون في أول الأمر بتهيئة المقالات التي يتلقونها من المؤلفين باستخدام تيجان أو وسيمات اللغة المعيارية الموحدة لتهيئة النصوص (SGML tags) ، والتي يتم بموجبها وصف بنية تلك المقالات ومحتواها. كما تقوم ما تعرف بنماذج عرض الصفحات أو الإخراج الطباعي Style Sheet بتحديد الكيفية التي يظهر بها كل عنصر من العناصر الهيكلية. وتعد كل من التهيئة باستخدام اللغة المعيارية الموحدة لتهيئة النصوص، ونماذج العرض أو الإخراج

الطباعي من مدخلات برمجية التحويل أو الترجمة التي تنشئ الوثيقة المنسقة formatted document.



الشكل رقم (9-1) العلاقة بين هيكل أو بنية الوثيقة ومظهرها

ومع بدايات ظهور المكتبات الرقمية أثير تساؤل عما إذا كانت هذه النوعية من المكتبات ستحل محل الكتب المطبوعة، كما تركزت المناقشات الأولية على قضية هامة هي مدى القدرة على القراءة readability، وتحت أي ظروف يمكن للناس أن يقرؤوا من شاشات الحاسبات بدلاً من الورق المطبوع؟. ومع خوض غمار التجربة أدرك الناس أن الحاسبات الإلكترونية والكتب المطبوعة لا يمكن أن يكونا

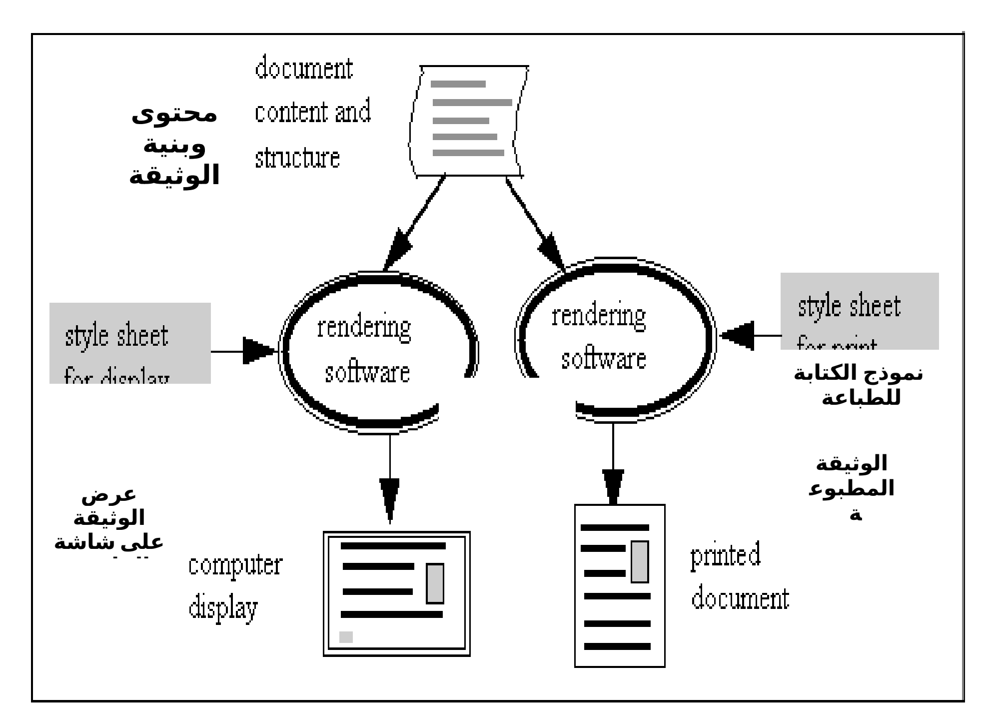
متكافئين تماماً، وإن لكل منهما مزايا لا تتوافر في الآخر؛ فالحاسبات الآلية تتوافر لها إمكانات بحث قوية لا يستطيع أي نظام يدوي الوصول إليها. كما أن للكتاب المطبوع من المزايا المرتبطة بالطابع البشري ما لا يتوافر للحاسبات؛ إذ يمكن على سبيل المثال حمله من مكان لآخر، كما يمكن تدوين بعض الملحوظات في هوامشه، هذا بالإضافة إلى إمكانية في أي وقت، أو على سطح المكتب أو حمله في إحدى اليدين، ناهيك عن عدم الحاجة إلى أي تجهيزات معينة لقراءته.

ولأن كلاً من النصوص الرقمية والمواد المطبوعة يخدم أغراضًا مختلفة، فقد عمد بعض الناشرين إلى إنتاج نسختين من المادة نفسها، إحداهما مطبوعة والأخرى إلكترونية متاحة مباشرة على الإنترنت. ويبين الشكل رقم (9-2) كيف يمكن استخدام أي من لغات التهيئة أو الترميز في إدارة النصوص التي ستتاح مطبوعة أو معروضة على شاشات الحاسبات. وباستخدام نماذج الكتابة أو عرض الصفحات المستقلة separate style sheet يمكن لوثيقة واحدة تم تمثيلها بواسطة التهيئة البنائية أن تحول بعدة طرق

ولأغراض مختلفة، حيث يمكن عرضها على الشاشة، أو طباعتها على ورق، مع ملاحظة أن إخراجها الطباعي قد يختلف عن تصميمها البنائي بالرغم من انبثاقهما عن مصدر واحد ويمثلان المحتوى نفسه.

ومن الصعوبة بمكان التحكم بشكل صارم في مظهر الوثائق عن طريق أي من أساليب الترميز سواء باستخدام اللغة المعيارية الموحدة لتهيئة النصوص (SGML) ونماذج الكتابة؛ لأن الطرق التي يتم بها تنظيم المواد النصية وعرضها وتحويلها تنطوي على كثير من الأمور الدقيقة. ويمكن للغات التميز أن تمثل جميع البنى تقريباً، مع أن تنوع العناصر البنائية في وثيقة ما قد يكون كبيراً، كما يمكن أن تكون تفاصيل المظهر التي يمكن للمؤلفين والمصممين اختيارها متفاوتة بالقدر نفسه، فعلى سبيل المثال تحتوي جميع مقالات الدوريات التي تنشرها جمعية صناع الحاسبات Association for computing Machinery على رموز رياضية، وغالباً ما يقدم المؤلفون مقالاتهم في صيغة تي إكس TeX التي توفر دعماً دقيقاً لإخراج الرموز الرياضية أو طباعتها. وأثناء عملية الإنتاج يتم تحويل هذه الصيغة إلى صيغة الترميز عن طريق اللغة المعيارية الموحدة لترميز النصوص (SGML)، وعندما يتم تحويلها باستخدام نماذج

الكتابة قد يبدو الشكل المخرج مختلفاً قليلاً عن شكل النسخة التي قدمها المؤلف.



الشكل رقم (9-2) التحويلات البديلة لوثيقة واحدة

تحويل النص :

بالرغم من أن الغالبية العظمى من الوثائق يتم إنشاؤها اليوم بشكل رقمي عن طريق الحاسبات، فإن المكتبات تزال مليئة بالوثائق القيمة التي لا تتوافر إلا في الورقية. ونتيجة لذلك فإن ثمة حاجة لتحويل المطبوعة إلى صيغ آلية، ومن ثم فإن أي مشروع الوثائق الهامة عليه أن يتعامل مع قضية المظهر ستبدو عليه هذه الوثائق، والتعرف إلى البنى الأصلية الوثائق.

ويعد المسح الضوئي هو الطريقة الأساسية للتحويل، حيث يتم مسح الوثيقة ضوئيًا عن طريق تعيين صورتها على شبكة مربعة من النقاط، وكل نقطة من هذه النقاط تمثيلها برمز يسمى رمز درجة الوضوح brightness code، وفي أبسط الأشكال لا يمكن إلا التمييز بين اللونين والأسود فقط. ويمكن أخذ صور جيدة لمعظم الصفحات المطبوعة عند درجة نقاء تصل إلى 300 نقطة في (سواء أكانت النقاط أفقية أو رأسية)، وإذا تم زيادة النقاء إلى 600 نقطة في البوصة، أو إذا تم ترميز مستويات من اللون الرمادي، فدرجة نقاء الصورة تصبح ممتازة، كما أن تناغم درجة الوضوح يمكن تمثيله وعرضه. ويحتاج العمل الفني عالي الجودة إلى 24 بت على في النقطة الواحدة حتى يمكن تمثيل التوليفات اللونية، وهذا يؤدي إلى إنشاء ملفات كبيرة جداً، وهذه الملفات ضغطها حتى تتلاءم مع عمليات التخزين والمعالجة، ذلك فإن أبسط الملفات النصية غير الملونة تحتاج 000 بايت على الأقل لتخزين صفحة واحدة منها فقط.

وتماثل الصفحة التي تم مسحها ضوئياً مظهر الصفحة المطبوعة، غير أن النص فيها يعرض كصورة لا أكثر. وفي العديد من التطبيقات تكون هذه الصورة بديلاً غير موفقٍ لنص تم تهيئته أو ترميزه marked up، أو حتى مجرد حروف بسيطة من رموز نظام الترميز القياسي الأمريكي لتبادل المعلومات المعروف بنظام آسكي ASCII.، وبشكل محدد، ليس من الممكن البحث عن كلمات محددة في صورة الصفحة، مع أنه في بعض الأحيان قد يحتاج بعض الباحثين إلى التعامل مع النسخ الأصلية، وغالبًا ما يكون في حاجة إلى معرفة الشكل الذي تبدو فيه الوثائق الأصلية. وفي حالات أخرى تكون النسخة الإلكترونية - التي تبين هوية بنية الوثيقة - أفضل؛ فعلى سبيل المثال: من الممكن أن يكون النص الإلكتروني المهيأ أو المرمز أكثر ملاءمة من صورة الأصل، وذلك عند عمل كشاف للألفاظ الواردة في النص Concordance، أو لأغراض التحليل النصي، ذلك لأن النص يمكن تتويجه tagged أو وضع تيجان لتدل على بنائه اللغوي أو على سياقاته التاريخية السابقة، ولذلك فإن المرحلة التالية من عملية التحويل تتمثل في توفير نص إلكتروني من صورة الصفحة.

ويعد التعرف إلى الأحرف بصرياً (Optical Characters Recognition OCR) هو الأسلوب الفني لتحويل صور الحروف الممسوحة ضوئياً إلى حروفها المقابلة لها، والأسلوب الفني الأساسي في هذه العملية هو عبارة عن برنامج آلي يقوم بفصل الأحرف الفردية، ثم مطابقة كل حرف من هذه الأحرف باستخدام نماذج أو برامج رياضية صغيرة mathematical templates، وبالرغم من مضي عـدة عقود من الأبحاث في هذا الصدد، فإن عملية التعرف إلى الأحرف بصرياً لا تزال غير دقيقة؛ حيث يتفاوت معدل الخطأ بمدى صلاحية النص الأصلي للقراءة، فإذا كانت الوثيقة الأصلية واضحة وقابلة للقراءة بشكل واضح، فإن معدل الخطأ يمكن أن يقل عن 1% ، أما إذا كانت النصوص الأصلية غير واضحة، فإن معدل الخطأ قد يرتفع إلى أكثر من ذلك بكثير. وفي حالات كثيرة يعد الخطأ كبيراً حتى لو كان كسراً في المائة لأن ذلك يؤدي إلى إيجاد حروف كثيرة غير واضحة في كل صفحة.

وقد تم ابتكار أساليب متنوعة للتغلب على هذه الأخطاء، ويتمثل أحد هذه الأساليب في استخدام عدة برامج مختلفة للتعرف إلى الحروف بصرياً للمواد نفسها التي يتم مسحها، بهدف إمكانية تجاوز الصعوبة التي تواجه أحد البرنامج وخاصة في تعرفها إلى بعض الحروف، عن طريق برامج أخرى. أما الأسلوب الآخر فيتمثل في استخدام المعجم لتصحيح نتائج المسح أو تفحصها. وبالرغم من ذلك فإن التدخل البشري لقراءة المسودات وتصحيحها أمر لا مفر منه للحصول على عملية تحويل عالية الجودة. وفي بعض النظم يقوم البرنامج الآلي بعرض النص الذي تم تحويله على الشاشة، مع تظليل الكلمات المشكوك فيها، وتقديم اقتراحات بالبدائل الصحيحة، وعلى المحرر أن يقبل هذه الاقتراحات أو يصحح ما يحتاج إلى تصحيح. وتعد شركة يو إم آي UMI التي تضطلع بمهام تحويل أعداد هائلة من الرسائل الجامعية في كل عام، من الشركات التي طورت عمليات ذات كفاءة عالية من هذا القبيل؛ ولأن الغالبية العظمى من الرسائل الجامعية التي تتلقاها هذه الشركة تكون في نسخ جديدة ونظيفة، فإن هذه الشركة تنجز عملها بمعدلات أخطاء منخفضة جداً عن طريق استخدام أسلوب التعرف إلى الأحرف بصريًا مصحوب بعملية تصحيح

يدوية للأخطاء في الكلمات المشكوك في صحتها.

وعندما يتم التعرف إلى الكلمات الفردية، تأتي الخطوة التالية من عملية التحويل، وهي تحديد بنية الوثيقة أو عناصرها البنائية، ثم تتويج عناوينها الرئيسية وغيرها من العناصر البنائية الأخرى في الوثيقة. وعلى الرغم من التقدم المستمر الذي يحرزه الباحثون في هذا الصدد، إلا أن التدخل البشري أمر لا مناص منه في عملية مراجعة التجارب الطباعية وتصحيح الأخطاء وفي عملية التحرير كذلك.

وثمة طريقة بديلة للتحويل، وهي في الحقيقة طريقة تستخدم ميدانياً على نطاق واسع، وتتمثل هذه الطريقة في إعادة كتابة الوثيقة كلها من مسودتها، وإضافة تيجان التهيئة أو الترميز يدوياً. وهذه الطريقة غالباً ما تكون أقل تكلفة من العملية المركبة التي تجمع بين المعالجة الآلية واليدوية، وحيث إن هذه الطريقة تحتاج إلى عمالة كبيرة فإنه عادة ما يتم تنفيذها في البلدان التي تكون فيها أجور العمالة منخفضة. ويعد برنامج الذاكرة الأمريكية American

memory - الذي تتبناه مكتبة الكونجرس- من أكبر مشروعات التحويل، حيث يتم اختيار الوثائق التي سيجري تحويلها من المجموعات التاريخية بالمكتبة، والتي تكون في العادة أقل وضوحاً من الوثائق حديثة الطباعة، وعادة ما يضطلع بمهام عمليات التحويل هذه مقاولون يتكفلون بضمان توفير قدر معين من الدقة، لكن لهم حرية اختيار الطريقة الملائمة التي ينفذون بها عملية التحويل. والجدير بالذكـر أن جميع المقاولين الذين اضطلعوا بهذه العملية في أول عهدها أقروا بأن طريقة إعادة كتابة الوثائق كانت من أكثر الطرق ملائمةً من الناحية الاقتصادية.

اللوحة رقم (9-1)

قاموس أكسفورد للغة الإنجليزية

تعد الطبعة الثانية من قاموس أكسفورد والتي نشرت في عام 1989م، خير مثال على استخدام إحدى لغات الترميز في وصف بنية النصوص ومحتواها. لأن المادة نفسها يمكن استخدامها كأساس لعدة أنواع مختلفة من المنتجات.

وقد أعد جيمس موريس James Morris وزملاؤه الطبعة الأولى من هذا القاموس على مدى أربعة عقود، وتظهر صور من القرن التاسع عشر لموريس وهو يعمل في مكان يحمل اسماً عظيماً هو حجرة النساخ Scriptorium في أكسفورد بأنه

في كوخ بدائي مليء بقصاصات الورق. وبالرغم من أن هذا القاموس كان علامة فارقة في تطور علم المعاجم، إلا أنه موجود فقط في صفحات مطبوعة ثابتة، وقد ثبت أنه من المستحيل تحديثه بطريقة يدوية.

وتمثلت الخطوة الأولى في إنتاج الطبعة الجديدة من هذا العمل في كتابة النص كاملاً من طبعته الأصلية وإدخاله إلى قاعدة بيانات، مع تحديد جميع الخصائص الطباعية المميزة، هذا فضلاً عن تحميل المادة الطباعية للقاموس بمعلومات عن معاني الكلمات، كما أن الحروف الكبيرة والسوداء الثقيلة والمائلة والأقواس والأشكال الصغيرة وغيرها من أساليب التنسيق الأخرى المستخدمة في الطباعة كانت أمورًا دقيقة لم تكن مسجلة في أي مكان آخر. وقد أعد برنامج آلي متطور يسعى إلى استخراج هذه المعلومات المتصلة بمعاني الكلمات، وكذلك إلى ترميز العناصر النصية textual elements عن طريق استخدام تيجان اللغة العامة الموحدة لترميز النصوص الفائقة (SGML). وقد تطلب هذا الجهد تكريس فريق عمل مكون من المتخصصين في لغات برمجة الحاسب الآلي Computational

Linguists في جامعة واترلو water looo في أونتاريو Onterio، بالإضافة إلى دعم من شركة آي بي أم (IBM) .

يحتفظ بقاموس أكسفورد للغة الإنجليزية في ثوبه الجديد هذا كقاعدة بيانات آلية، والتي تكتسي فيها عملية الترميز باستخدام اللغة المعيارية العامة الموحدة للترميز أهمية خاصة كأهمية كلمات هذا القاموس. ويقوم المعجميون بتحديث هذه القاعدة بشكل دوري، ويمكن من خلال هذه القاعدة - وبأقل جهد- استحداث إنتاج أشكال متنوعة من المخرجات، مثل الأسطوانات المدمجة، والإصدارات الرقمية، والكتب المطبوعة.

ترميز الحروف Encoding Characters :

الترميز المعياري الأمريكي لتبادل المعلومات (آسكي ASCII) :

تعد الحروف أهم العناصر الأساسية للنص، مثال ذلك الحرف (أ) والرقم (5)، ومن الأهمية بمكان التمييز بين مفهوم الحرف كعنصر بنائي، والأشكال المختلفة الممثلة لهذا الحرف والمختزنة في الحاسب الآلي أو المعروضة للقراءة؛ فالحرف هو مفهوم مجرد مستقل عن عملية الترميز المتبعة لاختزانه في الحاسب، وعن الشكل الذي يظهر به عند عرضه على الشاشة.

ويتم اختزان الحروف في الحاسبات كسلاسل متتابعة من البتات bits، وكل حرف مميز يكون مرمزاً كسلسلة متتابعة مختلفة عن غيرها من السلاسل. وكانت الحاسبات الآلية في مراحلها الأولى تحتوي على رموز لحروف اللغة الإنجليزية الست والعشرين (أحياناً ما تكون للحروف الكبيرة فقط)، وللأعداد العشرة، وللقليل من علامات الترقيم، ولبعض الرموز الخاصة، ولا يزال تمثيل الاختزان الداخلي في معظم الحاسبات الآلية معتمداً على مجموعة الحروف المحددة هذه. في حين تستخدم الغالبية العظمى من الحاسبات الحديثة الترميز المعياري الأمريكي لتبادل المعلومات المعروف بآسكي ASCII.

الجدول رقم (9-1)

الحروف القابلة للطباعة في نظام ترميز آسكي من فئة السبع بتات

32	Space	52	4	H	92	\	112	P
33	!	53	5	I	93	[	113	Q
34	"	54	6	J	94	^	114	R
35	#	55	7	K	95	_	115	S
36	$	56	8	L	96	??	116	T

37	%	57	9	M	97	a	117	U
38	&	58	:	N	98	b	118	V
39	'	59	;	O	99	c	119	W
40	)	60	>	P	100	d	120	X
41	(	61	=	Q	101	e	121	Y
42	*	62	<	R	102	f	122	Z
43	+	63	?	S	103	g	123	}
44	,	64	@	T	104	h	124	|
45	-	65	A	U	105	i	125	{
46	.	66	B	V	106	j	126	~
47	/	67	C	W	107	k	127	???
48	0	68	D	X	108	l		
49	1	69	E	Y	109	m		
50	2	70	F	Z	110	n		
51	3	71	G	]	111	o		

وقد قام نظام الترميز المعياري الأمريكي ASCII لتبادل المعلومات في الأساس بتمثيل كل حرف بسبعة بتات، ويعرف هذا الترميز ذو السبعة بتات بنظام آسكي القياسي Standard ASCII، فعلى سبيل المثال، يتم ترميز الحرف A في اللغة الإنجليزية بسلسلة متتابعة من سبعة بتات هكذا *1000001*، وعند النظر إلى هذه السلسلة على أنها رقم ثنائي فإنها تعني الرقم 65، ومن هنا فمن الطبيعي أن نقول عند التعامل مع ترميز آسكي القياسي ـ أن الرقم 65 يمثل الحرف A في الإنجليزية. وهناك 128 شكلاً مختلفًا يمكن

تركيبها عن طريق هذه البتات السبعة، ويعمل نظام آسكي القياسي على ربط حرف محدد بكل رقم من الأرقام الواقعة بين الصفر و 127، ومن هذه الأرقام، تخصص الأرقام من الصفر إلى 31 لتمثيل حروف التحكم [1] (مثل مفتاح العودة carriage return). ويبين الجدول رقم (9-1) رموز نظام آسكي التي تبدأ من الرقم 32 إلى 127، والتي تعرف بمجموعة حروف آسكي القابلة للطباعة (مع ملاحظة أن المسافة بين الحروف تحسب حرفاً من الحروف القابلة للطباعة).

وتعد مجموعة رموز آسكي القابلة للطباعة مجموعة قياسية بشكل حقيقي، حيث إن هذه الرموز نفسها تستخدم في أنواع كثيرة جداً من الحاسبات والتطبيقات. ولذلك فإن رموز آسكي الست والتسعين القابلة للطباعة تستخدم في تطبيقات يحظى فيها التشغيل المتداخل بأولوية كبرى، كما أنها الحروف الوحيدة المسموح باستخدامها مع لغة ترميز النصوص الفائقة (HTML) وفي العديد من نظم البريد الإلكتروني. كما أن جميع لوحات مفاتيح الحاسبات الآلية

---

[1]() يقصد رموز التحكم غير القابلة للطباعة (المترجمان).

وغيرها من أجهزة العرض وبرامج الحاسب الآلي تفسر هذه الرموز بالطريقة نفسها. وتجدر الإشارة إلى أن ثمة إصدارة موسعة من نظام رموز آسكي تستخدم نظام مكون من ثماني بتات eight bits [2]، وهي تقدم حروف ترميز إضافية للأرقام من 128إلى 255، ولكنها ليست مقبولة بالقدر الذي يقبل به نظام ترميز آسكي بالبتات السبعة seven-bits ASCII.

نظام الترميز الموحد (اليونيكود UNICODE) :

عادة ما تستخدم المواد النصية كما أكبر بكثير من الحروف التي تستخدمها مجموعة آسكي القابلة للطباعة، والتي لها أساسها في اللغة الإنجليزية. كما أن هناك بعض اللغات الأوربية لها حروف هجائية إضافية، أو أنها تستخدم علامات التشكيل مع بعض حروفها، بل إن اللغة الإنجليزية القديمة تحتاج إلى مزيد من الحروف الأخرى، أضف إلى ذلك أن لكل من اللغتين اليونانية والروسية حروفاً هجائية مختلفة، كما أن اللغات الكورية والصينية واليابانية تستخدم- فضلاً عن الحروف- رموز هان Han characters التي تمثل فيها كلمات أو مقاطع بأكملها برمز واحد فقط. وفي حقيقة الأمر ليست رموز آسكي كافية دائماً

2 (1) تعرف هذه الإصدارة بمجموعة الترميز الموسعة من آسكي Ascii Extended Characters (المترجمان).

لمتطلبات اللغة الإنجليزية الحالية أو المعاصرة. كما أن ثمة تخصصات علمية معينة (مثل الرياضيات و الموسيقى والكيمياء) تستخدم نظام ترميز فائق الدقة يتطلب أعداداً كبيرة من الحروف، ولا شك أن دعم مقومات القدرة على الفهم يعتمد أساساً على استخدام ضوابط مقبولة لنظام الترميز.

إن شركات تصنيع الحاسبات الآلية تبيع منتجاتها في جميع أنحاء العالم، وتدرك الحاجة إلى دعم الحروف التي يستخدمها عملاؤها حول العالم، وهذا هو المجال الذي حققت فيه الريادة شركة تتعرض للكثير من الحقد والعداوة، أعني بها شركة مايكروسوفت. ولأنه من غير الممكن تمثيل كل اللغات باستخدام الاحتمالات المائتين والست والخمسين (256) الممثلة بالبتات المعتمدة على نظام البتات الثمانية an eight-bit byte، كانت هناك عدة محاولات لتمثيل نطاق أكبر من مجموعات الرموز باستخدام عدد أكبر من البتات. وقد برزت إحدى هذه المحاولات في الآونة الأخيرة باعتبارها معياراً تدعمه معظم

شركات إنتاج الحاسبات والبرمجيات، ويعرف هذا المعيار باسم الترميز أو الكود الموحد (اليونيكود UNICODE).

في معيار اليونيكود الدقيق يمثل كل حرف بستة عشر بتاً (*16* bits)، وهذا يعني إمكانية تمثيل أكثر من 65536 حرفاً مميزاً. ويمكن الآن بفضل الجهود الجبارة التي اضطلع بها عدد من المختصين المخلصين في نظام اليونيكود تمثيل هجائيات عدد كبير من اللغات.

ومما تجدر الإشارة إليه أن تقبل نظام اليونيكود لا يعود فقط إلى جهود علماء اللغويات، بل إلى جهود المطورين الذين فكروا بعناية عن مدى العلاقة بين نظام اليونيكود والبرمجيات المتوافرة. ولو كان استخدام هذا النظام يستلزم تغييراً لكل برامج الحاسبات المستخدمة لما تبناه أحد. لذلك هناك نظام تمثيل خاص لرموز اليونيكود يعرف بيوتي إف ـ 8 (UTF-8)، يسمح بتحويل التطبيقات المعتمدة على نظام آسكي تدريجياً إلى نصوص اليونيكود unicode scripts بشكل كامل.

ويو تي إف - 8، هو نظام ترميز يستخدم ما بين 1-6 بايتات لتمثيل كل حرف من حروف اليونيكود، بحيث تمثل أكثر الحروف شيوعاً واستخداماً ببايت واحد، في حين تمثل الحروف الأقل شيوعاً منها ببايتين، ثم الأقل منها شيوعاً

بستة بايتات. والجانب الأساسي في هذا التصميم أن كل حرف من حروف آسكي القابلة للطباعة يمثل ببايت واحد، وهو مطابق بذلك لنظام رموز آسكي المناظر له. ومن ثم يمكن تفسير السلسلة المتتابعة نفسها من البايتات على أنها إما رموز يونيكود (في تمثيل يوتي إف ـ 8) أو رموز آسكي القابلة للطباعة. فعلى سبيل المثال، إذا كانت هناك صفحة بها نص مرمز باستخدام لغة ترميز النصوص الفائقة (HTML) وتم استحداث هذه الصفحة باستخدام رموز آسكي القابلة للطباعة، فإن هذه الرموز لا تحتاج إلى أي تعديلات لاستخدامها في برنامج يتوقع أن تكون بياناته مكتوبة بنظام الترميز يوتي إف - 8.

اللوحة رقم (9-2)

الخطوط الممثلة في نظام اليونيكود

تشتمل الإصدارة التي تحمل رقم version 2.0 من مواصفة اليونيكود على رموز من فئة الستة عشر بتاً (*16* bits) تمثل 38885 حرفاً متميزاً، ويتم تنظيم عملية الترميز على أساس الخطوط أو الهجائيات scripts، وليس على أساس

اللغات، فحيثما تكون هناك عدة لغات تستخدم المجموعة نفسها من الحروف المرتبطة مع بعض ارتباطاً وثيقاً، فإن مجموعة الرموز symbols التي تغطي مجموعة من اللغات تعرف على أنها خط واحد. فالخط اللاتيني - على سبيل المثال - يضم كل الحروف المستخدمة في اللغات الإنجليزية والفرنسية والأسبانية والألمانية واللغات القريبة منها، كما أن إصدارة يونيكود UNICODE تدعم اللغات العربية والأرمينية والبنغالية والبوبوموفو Bopomofo والسيريلية، والديفنجارية Devangari والجورجية واليونانية والقوجاراتية Gujarati والجورمخية Gurmkhi ولغة الهان Han والهانجول Hangul والعبرية والهيراجانا Hiragana والكنادا Kannada والكاتاكانا Katakan ولغة اللاو Lao والملايالان Malayalan والأوريا Oriya والفونتيك Phonetic والتاميلية والتليوجو Tolugu والتايلندية ولغة التبت Tibet.

وفضلاً عن تلك الخطوط الرئيسية المشار إليها في الفقرة السابقة هناك مجموعات أخرى من الرموز التي تمثل الأرقام وعلامات التشكيل العامة، وعلامات الترقيم العامة، والرموز العامة، والرموز الرياضية، والرموز الفنية، وعلامات الزخرفة Dingbat، والعلامات المطبعية في شكل الأسهم arrows، والكليشيهات، وأشكال الرسم، والأشكال الهندسية، وأشكال العرض presentation forms. وسوف يتم دعم العديد من

اللغات الأخرى حديثة النشأة (مثل الأثيوبية والسنهالية)، كما سيتم في المستقبل دعم العديد من الخطوط الميتة أو القديمة (مثل الآرامية Armaic، والإتروسكانية Etruscan، والريونز Runes).

وهناك مظهر مهم وخاص بنظام يونيكود وهو أنه يدعم حروف هان المستخدمة في اللغة الصينية واليابانية والكورية. كما أنه نظام متوافق مع يوني هان Unihan، الذي كان نتيجة لمشروع مبكر لإصلاح نظام الترميز الذي كان مستخدماً في السابق لهذه اللغات.

النقل الصوتي للحروف Transliteration :

ليس نظام اليونيكود هو الوسيلة الوحيدة المستخدمة لتمثيل طائفة كبيرة من الحروف في الحاسب الآلي، وإنما هناك طريقة أخرى هي طريقة النقل الصوتي للحروف، وهي طريقة منتظمة لتحويل الحروف من هجائية لغة معينة إلى مجموعة أخرى من الحروف، فعلى سبيل المثال، الحرف "ö" في اللغة الألمانية ينقل صوتياً في غالب الأحيان إلى " oe". وهناك نظام آخر للنقل الصوتي يعرف

باسم بنين pinyin، كثيراً ما يستخدم لتمثيل حروف اللغة الصينية (وخاصة الماندرينية Mandarin) بحروف اللغة الإنجليزية. وربما كان نظام النقل الصوتي للحروف مقبولاً عندما كانت الآلات الكاتبة تعمل بطريقة يدوية، وما كان يصاحب ذلك من صعوبات كثيرة في كتابة مجموعة من الحروف. ولكن مع استخدام الحاسبات لا ينبغي أن تكون هناك حاجة لنظام النقل الصوتي للحروف، وحسبه أنه سيصبح في المستقبل القريب من أطلال الماضي.

ولقد كانت المكتبات تستخدم مجموعة كبيرة من الهجائيات ولفترة طويلة قبل أن تولي صناعة الحاسبات اهتمامها بالخطوط، وقد حققت المكتبات في واقع الأمر تقدماً كبيراً في هذا الجانب، في الوقت الذي كانت فيه معظم الحاسبات تدعم الحروف العالية أو الاستهلالية فقط upper-case. ونتيجة لذلك، فإن فهارس مارك المقروءة آلياً، وغيرها من نظم المكتبات تحتوي على مجلدات هائلة من المواد التي تم ترميزها بنظام بنين (Pinyin) للنقل الصوتي للحروف، بغيرها من النظم الأخرى بالإضافة إلى نظام اليونيكود، وتواجه المكتبات الآن صعوبات كثيرة سواء عند إجراء عمليات التحويل، أو عند محاولة التعايش مع كل هذه النظم المختلفة.

اللغة المعيارية الموحدة لترميز النصوص أو

تهيئتها (SGML) :

استخدمت لغات الترميز أو التهيئة منذ بداية استخدام الحاسبات الآلية لوصف بنية النصوص، ولوصف الأشكال أو الصيغ التي ستعرض بها تلك النصوص على شاشات الحاسبات أمام القارئ، ومن أكثر لغات الترميز شيوعاً من حيث تواتر استخدامها هي تلك التي من عائلة اللغة المعيارية الموحدة لترميز النصوص (SGML).

واللغة المعيارية الموحدة لترميز النصوص SGML ليست لغة ترميز واحدة، وإنما هي نظام يستخدم لتعريف محددات الترميز markup specifications، وأي محدد فردي يتم تحديده في إطار اللغة المعيارية الموحدة لترميز النصوص تسمى معرف نوع الوثيقة (Document Type Definition (DTD)).

وقد قام عدد من الناشرين، وبعض الجهات الأخرى ذات الصلة بهذا المجال، بتطوير " معرفات أنواع الوثائق DTDs" الخاصة بهم، وثمة معرفان لأنواع الوثائق يكتسبان أهمية خاصة بالنسبة للمكتبات البحثية، وهما: "مبادرة ترميز النصوص Text Encoding Initiative (TEI) ، والوصف الأرشيفي

المرمز Encoded Archival Description (EAD) ، وتتضمن اللوحة رقم (9-3) وصفاً لهذين المعرفين.

ويعتمد معرف نوع الوثيقة في إنشائه على المفاهيم العامة لكل من الكائنات entities والعناصر elements، وهو يهدف إلى تعريف الكائنات والعناصر المسموح بها في فئة محددة من الوثائق، كما أنه يعمل على توضيح النظام الأساسي لترميز مجموعة الحروف المستخدم لهذه الوثيقة.

وتحدد الكائنات عن طريق المحددات التي تبدأ عادة بعلامة (&)، وتنتهي بعلامة الفاصلة المنقوطة (;)، هكذا على سبيل المثال :

α

&logo;

وفي معرف نوع الوثيقة المثالي، تشتمل الكائنات المسموح بها على معظم حروف نظام آسكي، وعلى بعض الحروف الأخرى (والتي تعرف بكائنات الحروف character entities، ويمكن - على أية حال - أن يُعرّف أي رمز symbols، أو مجموعة رموز على أنه كائن واحد. أما اسم الكائن فهو مجرد اسم. وفي مجموعات الحروف القياسية، يعد الكائن الموسوم بـ "α"، هو الكائن المستخدم لترميز الحرف

الأول من الهجائية اليونانية، ومع ذلك، فإن معرف نوع الوثيقة يمكنه استخدام هذا الرمز لغرض آخر مختلف تماماً. وتحدد معرفات أنواع الوثائق التي يستخدمها الناشرون العلميون ما يزيد على 4000 كائن مستقل لتمثيل جميع الرموز الخاصة وأشكالها المختلفة المتداولة في التخصصات العلمية.

وتوفر الكائنات سيلاً من الرموز التي يمكن تجميعها معاً في شكل عناصر، كما يستطيع معرف نوع الوثيقة أن يحدد أي سلسلة متتابعة من الرموز على أنها اسم لعنصر معين، والعنصر عادة ما يتم حصره بين وسيمتين أو تاجين في شكل الأقواس ذات الزوايا (< >)، مع استخدام الشرطة المائلة مع التاج الأخير (/>) لتبين نهاية هذا العنصر أو حدوده، ولذلك فإن مبادرة تر ميز النص تستخدم التاجين التالين، وعلى هذا النحو: "<del> </del>"، لحصر نص يتم حذفه من المخطوطة الأصلية التي ورد فيها، فعلى سبيل المثال، عند الإشارة إلى حذف جملة "men and women" من نص ما، يتم تتويجها على النحو التالي: <del> men </del> and women

وتضم قائمة العناصر أنواعًا كثيرة مثل: العناوين الرئيسية، والحواشي، وقائمة المصطلحات، وقائمة المراجع. ويمكن تجميع العناصر معاً في شكل علاقات هرمية، وأخيراً تجدر الإشارة إلى أن لكل نوع من معرفات أنواع الوثائق قواعده التي تحدد العلاقات المسموح بها كمجموعة من القواعد التي يمكن معالجتها عن طريق البرنامج الآلي.

اللوحة رقم (9- 3)

معرفات أنواع الوثائق **DTDs** للدراسة العلمية

كانت مبادرة ترميز النص من الجهود المبكرة والشاملة لتمثيل النصوص الموجودة في أشكال رقمية، مع التركيز على النصوص التي تتصل باحتياجات الباحثين في مجال الإنسانيات. فالنصوص القديمة تعرض في كل الوسائل التي سبق أن استخدمت في تدوين النصوص سواء على الورق أو على الجلود أو على ورق البردي أو حتى على الحجارة، وسواء تم هذا التدوين بأساليب الطباعة أو بالكتابة بخط اليد أو بالرقم أو بالآلات الكاتبة أو غير ذلك، وكثيراً ما تشتمل الوثائق القديمة على شروح أو تفسيرات و أجزاء محذوفة و أجزاء تالفة لا يمكن قراءتها، ومن الممكن كذلك توقع وجود أي مجموعة حروف لم يسبق أن استخدمت من

قبل هذه الوثائق.

وقد أثبتت اللغة المعيارية الموحدة لترميز النصوص كفاءة في وصف هذه المواد [القديمة]، غير أن تصميم معرف نوع الوثيقة الخاص بها كان تحدياً، وربما كان من السهولة بمكان أن تنشئ معرفاً شديد التعقيد بحيث يصعب تداولـه بشكل واسع. وكان الحل هو في استخدام مجموعة مكونات component لبناء مجموعة من معرفات أنواع الوثائق DTDs. ومن ثم فقد تم تضمين جميع معرفات أنواع الوثائق مجموعة التاج الأساسي a core tag set التي تحدد العناصر المرجح أن تكون مطلوبة لجميع أنواع الوثائق. كما أن لكل معرف من تلك المعرفات مجموعة تاج أساسية base set من التيجان التي أنشئت لتتناسب مع النصوص النثرية، والشعرية، والدرامية، والمعاجم، وملفات البيانات. وعادة ما تكون مجموعة أساس واحدة فقط ملائمة لوثيقة معينة. وأخيراً هناك مجموعة متنوعة من مجموعات التيجان الإضافية المتاحة لأغراض خاصة. (وتجدر الإشارة إلى أن المؤلفين يطلقون على هذه المجموعات نموذج بيتزا

شيكاجو Chicago pizza؛ لأن هذا النوع من البيتزا عادة ما يتألف من قطعة من الجبن وصلصة الطماطم، مع طبقة أساس واحدة، مع بعض الإضافات الاختيارية الأخرى التي يحددها المشتري.

أما مصطلح "وسائل البحث أو معيناته finding aids" فيشمل عدداً كبيراً من القوائم، والكشافات وغيرها من الوثائق النصية الأخرى التي تستحدثها المكتبات والأرشيفات والمتاحف لوصف مقتنياتها. وعادة ما تقدم معينات البحث معلومات أكثر من تلك التي تتضمنها تسجيلات الفهارس المعروفة، وإن كان بعضها قد يتضمن معلومات أقل تحديد، ولا يشتمل بالضرورة على تسجيلات تفصيلية لكل عنوان في المجموعة، كما أن بعض وسائل البحث هذه قد يكون مختصراً، في مقابل ذلك قد يأتي بعضها الآخر في مئات من الصفحات.

والوصف الأرشيفي المرمز Encoded Archival Description (EAD) هو عبارة عن معرف من معرفات أنواع الوثائق يستخدم لترميز الإصدارات الإلكترونية من معينات البحث الأرشيفية archival aids. وقد اضطلع بمهام تطوير الإصدارة الأولى من هذا المعرف فريق عمل من جامعة كاليفورنيا في بيركلي. وقد أفادوا من تجربة مبادرة ترميز النص Text

encoding initiative وتعكس نتيجة جهودهم في هذا العمل طبيعة المفاهيم النظرية والممارسات العملية التي أرساها الأرشيفيون، كما تعكس كذلك الطبيعة التركيبية المعقدة لمعينات البحث، وبالرغم من أن كثيراً من المعلومات تكون مستقاة من العلاقات الهرمية، فإن هناك علاقات أخرى كثيرة متداخلة يجب على المرء أن يعرفها صراحة عند ترميز معينات البحث لأغراض استخدامها في المكتبات الرقمية.

لقد لاقى الوصف الأرشيفي المرمز EAD قبولاً من كثير من الأرشيفيين الذين تضافرت جهودهم لتنقيحه واختباره في مقابل معينات البحث الأخرى الموجودة، كما أنهم تعاونوا في توثيق هذا العمل بشكل شامل، وقد بات هذا العمل أداة في غاية التخصص، ومفصلة لتلبية احتياجات المجتمع المتخصص عند تبادلهم وتقاسمهم للمعلومات.

الإصدارات المبسطة :

لقد باتت اللغة المعيارية الموحدة لترميز النصوص (SGML) طريقة راسخة ومرنة من طرق تسجيل النصوص

ذات الدقة العالية وتخزينها. وتتيح مرونة هذه اللغة لمنشئي المواد النصية أن ينشئوا معرفات لأنواع الوثائق مفصلة وفق احتياجاتهم. وفي اللوحتين (9-1)، (9-3) وصف لبعض الأمثلة على ذلك. ويذكر في هذا السياق أن ناشري الدوريات العلمية يعتادون على تطوير معرفات أنواع الوثائق الخاصة بهم والتي يستخدمونها في ترميز مقالات الدوريات عند اختزانها بالحاسبات. كما تستخدم مشروعات المكتبات الرقمية- كمشروعي جستور JSTOR، والذاكرة الأمريكية American memory- معرفات أنواع وثائق معتمدة في فكرتها على جهود "مبادرة ترميز النص TEI".

ولعل أهم ما يسجل من نقاط الضعف المرتبطة بمرونة اللغة المعيارية الموحدة لترميز النصوص SGML يتمثل في تعقد البرمجية المطلوبة لتجهيزها أو معالجتها. ومع أنه ليس من الصعوبة كتابة التحليل اللغوي وترجمة معرف بسيط من معرفات أنواع الوثائق، فإنها لمهمة صعبة أن تنشئ حزمة برمجية متعددة الأغراض تستطيع أن تحلل بنية أي معرف، أو أن تدمج معلومات من أي نموذج للكتابة style sheet أن تنقل الوثيقة إلى شاشة الحاسب أو إلى الطابعة. يضاف إلى ذلك أن سوق مثل هذه الحزمة البرمجية هي سوق صغيرة إلى حد ما. ومن الملاحظ أن ثمة شركة واحدة هي التي راهنت على إعداد هذه البرمجية، وقد

أعدت بالفعل حزمة برمجية تستخدم لأغراض تحويل أو ترجمة اللغة المعيارية الموحدة لترميز النصوص. ومع كون هذه الحزمة البرمجية ليست على الوجه الأكمل، فإنها لم تستطع تنفيذ كل ما في اللغة المعيارية العامة الموحدة لترميز النصوص، كما أنها تعمل مع بعض أنواع الحاسبات فقط، هذا فضلاً عن أنها تستخدم شكلاً خاصًا بها من نماذج الكتابة. ومن ثم فإن اللغة المعيارية الموحدة في صورتها الكاملة ليست ملائمة للاستخدام في المكتبات الرقمية التي تحرص في عملها على ضرورة التشغيل المتداخل لأنظمتها مع الأنظمة الأخرى.

لغة ترميز النصوص الفائقة (HTML) :

لقد شجعت الويب على تطوير إصدارات مبسطة من اللغة المعيارية الموحدة لترميز النصوص (SGML). وتعد لغة ترميز النصوص الفائقة (HTML) المستخدمة مع الويب بمثابة معرف تقليدي من معرفات أنواع الوثائق، ومع ذلك فهي تختلف في فلسفتها عن فلسفة اللغة المعيارية الموحدة (SGML)، حيث تسعى إلى دمج المعلومات البنائية مع المظهر أو الشكل. وقد انصب الهدف الأساسي من لغة ترميز

النصوص الفائقة في أول الأمر على استخدامها لأغراض الترميز البنائي معتمدة في ذلك على المتصفحات لتحدد كيفية تنسيق النص للعرض، غير أن تطورها اللاحق أضاف عدداً كبيراً من الخصائص التي تمنح مصمم صفحات الويب القدرة على التحكم في مظهر المادة عندما يتم نقلها للعرض على الشاشة أو عند طباعتها. وتبين اللوحة رقم (9-4) مدى اختلاف لغة ترميز النصوص الفائقة عن عملية الترميز البنائي البحت.

وتواصل لغة ترميز النصوص الفائقة نموها وتطورها وإن كان هذا التطور وما صاحبه من إضافات لخصائص جديدة يوصف من وجهات نظر مختلفة على أنه يهدف إلى تقوية هذه اللغة، في حين ترى وجهات نظر أخرى أن هذه الإضافات ستضفي على هذه اللغة المزيد من التعقيد وعدم الملاءمة. وفي الوقت ذاته لا يمكن لأحد أن ينكر القيمة الكبرى لبعض الإضافات، كما هو الحال بالنسبة لخاصية إمكانية إدراج الصور من خلال التاج <image> (الذي قُدم في برنامج موزايك mosaic)، أما قيمة الإضافات الأخرى فإنها تخضع لاختلاف وجهات النظر، فأوامر التنسيق formatting البسيطة مثل الأمر <center> (الذي قدمه متصفح نيتسكيب Netscape) قد يؤدي إلى بعض الأضرار القليلة، في حين أضافت بعض الخصائص الأخرى قدراً كبيراً من التعقيد.

وربما كان من أبرز تلك الخصائص ما يتصل منها بإعداد الجداول والإطارات. وعليه فإنه يمكن القول بأنه لم يعد باستطاعة أي مؤلف أن يتعلم لغة ترميز النصوص الفائقة بين عشية وضحاها، كما لم يعد باستطاعة أي مبرمج أن يكتب برنامجاً لترجمة هذه اللغة أو تفسيرها في أسبوع واحد.

لقد تفاقمت مظاهر التوتر بين الترميز البنائي في لغة ترميز النصوص الفائقة ومساعي التنسيق أو التشكيل من أجل التحكم في المظهر، حيث يسعى كثير من منشئي صفحات الويب، انطلاقاً من رغبتهم في التحكم وضبط ما يراه المستفيد، إلى انتهاز كل فرصة تقدمها هذه اللغة للتحكم في المظهر، وهناك توجه غير موفق يتمثل في قيام المصممين باستخدام العناصر البنائية في معالجة التصميم، وقد ُصممت كثير من صفحات الويب على هيئة جدول واحد ضخم، بحيث يمكن للمصمم أن يتحكم في ضبط الهوامش التي يراها المستفيد، ويقوم المصممون المهرة بإنشاء صفحات ويب رائعة باستخدام مثل هذه الحيل، أما المصممون الأقل مهارة فيقومون بتصميم

صفحات غير مقبولة للاستخدام، ولعل من أسوأ أشكال الاستخدام، هو ما يحدث عندما يقوم بعض المصممين بتصميم تخطيط صفحة لا تتناسب وحجم النافذة التي يستخدمها المستفيد، أو لا تتفق والخيارات التي يريدها، بل قد يتطور الأمر في بعض الحالات إلى أن هذا التصميم السيئ قد يحرم المستفيدين ضعفاء النظر من استخدام الحجم الأكبر.

وقد جاءت جميع الخصائص تقريباً التي أضيفت إلى لغة ترميز النصوص الفائقة من قبل مطوري المتصفحات الذين عملوا على إضافة خصائص لتحسين منتجاتهم أو لمجاراة منافسيهم، وإذا استطعنا أن نصف بعض هذه الخصائص على أنها من قبيل التحسينات، فإنه يمكننا القول بأن ثمة خصائص ما هي إلا ترف، ويقوم كل من اتحاد الويب www consortia ولجنة هندسة الإنترنت internet engineering task بالتنسيق فيما بينهم بصورة قيمة، كما يحاولون كذلك تقديم معايير في هذا الصدد، ولكن لا ننسى أن التحكم الأساسي لمجريات الأمور في لغة ترميز النصوص الفائقة يمارسه المصنعان الأساسيان للمتصفحات، وهما شركتا : مايكروسوفت ونيتسكيب.

اللوحة رقم (9-4)

خصائص لغة تهيئة أو ترميز النصوص الفائقة أو ترميزها (
(THML

تظهر معظم تيجان لغة ترميز النصوص الفائقة على شكل زوجين من التيجان، فعلى سبيل المثال، يشير هذا الزوج من التيجان <ht> text </ht> إلى أن النص المحصور بينهما هو عنوان رئيسي من المستوى الأول، وهناك تيجان قليلة تأتي في زوج واحد ولا تحصر أي نص بينها، ومثال ذلك التاج <hr> الذي يشير وضع خط أفقي horizontal rule بين فقرات النص، وأحياناً ما يتم تجميع العناصر معاً أو حصرها بعضها مع بعض، كأن تحتوي قائمة على فقرات كثيرة، وقد يحتوي جدول على جداول أخرى.

والأمثلة التالية هي نماذج للخصائص التي توفرها لغة ترميز النصوص الفائقة، وهي تبين دمج هذه اللغة عملية التهيئة البنائية مع عملية التنسيق أو التشكيل، وكيفية دعمها لأغراض التطبيقات المباشرة (وهناك العديد من الخصائص الأخرى، كما أن لكثير من هذه الخصائص المبينة في الفقرات التالية مجالات واسعة من الخيارات). وفيما يلي

بعض التيجان المستخدمة لوصف العناصر البنائية :

- التاج: <body>، الذي يشير إلى جسم الوثيقة.
- التاج: <p>، الذي يشير إلى الفقرة.
- التيجان: <h1>,<h2>,....<h6>، التي تشير إلى أن هناك رؤوس موضوعات، ترد في ستة أحجام قياسية.
- التاج : <em>، الذي يشير إلى التركيز.
- التيجان : <ul>,<ol>,<dl>، التي تشير إلى القائمة غير مرتبة، ثم القائمة مرتبة، ثم قائمة التعريفات.
- التاج : <table>، الذي يشير إلى وجود جدول.

كما تشمل التيجان التي تحدد مظهر الوثيقة عند إرسالها للعرض على شاشة الحاسب الآلي أو عند طباعتها ما يلي:

- التاج:
، الذي يشير إلى انتهاء السطر أو الفقرة.
- التاج <i>، الذي يشير إلى أن الخط يجب أن يكتب مائلاً.
- التاج <font> الذي يشير إلى بيانات عن حجم الخط الذي سيتم استخدامه.
- التاج <center>، الذي يشير إلى توسيط النص.
- التاج <pre>، الذي يشير إلى نص سبق إعداده من

قبل.

كما تضم التيجان المصممة للتطبيقات المباشرة ما يلي:

- التاج <a>، الذي يشير إلى المرتكزات anchor المستخدمة مع الروابط الفائقة.
- التاج <img>، الذي يشير إلى وضع صورة.
- التاج <form>، الذي يشير إلى شكل متاح مباشرة أو على الخط المباشر online form.

لغة التهيئة أو الترميز الموسعة (XML) :

لغة الترميز الموسعة (XML) هي شكل آخر مختلف من اللغة المعيارية الموحدة لترميز النصوص (SGML)، وهي تسعى إلى سد الفجوة الفاصلة بين بساطة لغة ترميز النصوص الفائقة (HTML) وقوة اللغة المعيارية الموحدة للترميز (SGML)، وإذا كانت البساطة هي سر نجاح لغة ترميز النصوص الفائقة فإنها تعد في الوقت ذاته من نقاط الضعف التي اعتورتها. ففي كل مرة أضيفت فيها خاصية إلى تلك اللغة كان لها أثرها السلبي الواضح على جمال هذه اللغة وعلى استخدامها ومعاييرها مقارنة بالمستوى الذي تشترك

فيه جميع المتصفحات. وفي المقابل تتمتع اللغة المعيارية الموحدة بمرونة أكبر تجعل إمكانية الوصف النصي text description أمراً ممكناً في معظم الأحوال، غير أن هذه المرونة عادة ما تكون على حساب مستوى تعقدها.
وبالرغم من مرور سنوات عديدة على تطوير هذه اللغة، فإن النزر اليسير فقط من المتخصصين هم من يأنسون إلى استخدام هذه اللغة، هذا فضلاً عن أن واقعها لا يزال يفتقر - إلا فيما ندر - إلى برمجية من البرمجيات التي تعرف ببرمجيات الأغراض العامة.

ولغة الترميز الموسعة (XML) بمثابة مجموعة فرعية من اللغة المعيارية الموحدة للترميز (SGML)، تم تطويرها بناءً على أساسين هامين، يتمثل أولهما في التزام البساطة في كتابة برامج آلية لمعالجتها، أي لغة الترميز الموسعة (XML)، ويتمثل ثانيهما في تمكين كل من الأفراد والنظم - وبأقل قدر من العنت - من الانتقال من لغة ترميز النصوص الفائقة إلى لغة الترميز الموسعة.

والكود الموحد من فئة الستة عشر بتاً 16-bit code وتحديداً نظام الترميز المسمى يو تي إف ـ 8 هو الذي يشكل مجموعة الحروف الأساسية للغة الترميز الموسعة. وهذا الأسلوب من أساليب الترميز يسمح بكتابة الوثائق بنظام

آسكي ASCII القياسي، كما أنه يدعم مجموعة كبيرة من اللغات ومجموعات الحروف. وفي مساعيها لتوفير مقومات الملاءمة أو السهولة، تم تحديد بعض كائنات الحروف تحديداً مسبقاً، مثال ذلك الرمزان : " lt &"، و "gt &" ، اللذان يرمزان لكلمتي: "أقل من"، و "أكثر من". وتجدر الإشارة إلى أن لغة الترميز الموسعة لا تحدد أساليب معينة لتمثيل الرموز الرياضية، ولكن هناك جهوداً مستقلة تعرف بلغة ترميز النصوص الرياضية "ماث إم إل math ML" تعمل على تلبية هذه الحاجة.

إن لغة ترميز النصوص الفائقة القياسية standard HTML، وما صاحبها من تعديلات طفيفة، يمكن أن تقبل كلغة الترميز الموسعة (XML)، وقد تمثل أحد هذه التعديلات في أن تيجان النهاية دائماً ما تكون مطلوبة ولا يمكن الاستغناء عنها، فعلى سبيل المثال، إذا كان التاجان <p> </p> في لغة ترميز النصوص الفائقة يحددان بداية الفقرة ونهايتها، فإن التاج </p
>، هو تاج اختياري يستخدم عندما تكون هناك فقرة تالية لتلك الفقرة التي يتم تحديدها. وفي لغة الترميز الموسعة

دائماً ما يكون تاج النهاية مطلوباً ولا يمكن الاستغناء عنه، وهناك تعديل آخر كبير يتصل بتيجان لغة ترميز النصوص الفائقة التي لا تحدد أي محتوى. فعلى سبيل المثال يشير التاج
 إلى نهاية السطر في لغة الترميز الموسعة (XML) ، أما في لغة ترميز النصوص الفائقة القياسية فإن التاج
 لا يستخدم، وفي لغة الترميز الموسعة يتوج نهاية السطر إما بزوجي التيجان
</br> أو بتاج أقصر حيث توضع الشرطة المائلة بعد الرمز وليس قبله هكذا:
 .

وبما أن لغة الترميز الموسعة (XML) هي عبارة عن مجموعة فرعية من اللغة المعيارية الموحدة للترميز (SGML)، فإن كل وثيقة تعتمد في ترميزها على أحد معرفات أنواع الوثائق. غير أن هذا المعرف لنوع الوثيقة ليس من الضروري تحديده صراحة. وإذا كان الملف يشتمل على زوجين من التيجان غير محددين مسبقاً وهما يحددان بعض أجزاء الوثيقة، فإن المحلل الآلي parser يقوم بإضافتهما تلقائياً إلى معرف نوع الوثيقة.

لقد بذل مطورو لغة الترميز الموسعة جهوداً جبارة ليحظى عملهم هذا بالقبول الواسع، وقد سارت استراتيجيتهم في العمل وفقاً للفلسفة العامة للإنترنت. وقد كان التصميم لهذه اللغة بمثابة عملية مفتوحة

استضافها اتحاد الويب www consortium، وقد عمد فريق التصميم منذ البداية إلى كتابة برمجية عرض ووزعها مجاناً على جميع الأطراف المعنية بهذا المجال، وخاصة شركتي ميكروسوفت ونيتسكيب، اللتين قدمتا دعمهما للفريق، ويبدو أن هذا الحشد القوي والدعم آتيا ثمارهما، إذ يبدو من المرجح أن يتم تبني لغة الترميز الموسعة على نطاق واسع.

نماذج الكتابة أو عرض الصفحات Style Sheets :

تصف لغات الترميز بنية الوثائق، وتستخدم اللغة المعيارية الموحدة لترميز النصوص (SGML) ، ولغة الترميز الموسعة (XML) التيجان لوصف العناصر الدلالية semantic للوثيقة وعناصرها الأساسية الأخرى، ولكنهما لا تصفان شكل الوثيقة أو مظهرها، ومن ثم فإن تيجان اللغة المعيارية الموحدة يمكن استخدامها لتحديد جزء من النص كالعنوان الرئيسي أو عنوان الفصل، ولكنها لن تشير إلى أن عنوان الفصل هذا - مثلاً - يبدأ من صفحة جديدة، أو يطبع بحجم معين أو يتم طباعته بترتيب معين.

وهناك حاجة لا خلاف عليها تتمثل في أن أخذ الوثيقة

المرمزة بتيجان اللغة المعيارية الموحدة أو لغة الترميز الموسعة وترجمتها أو نقلها يجب أن يكون وفقاً لمجموعة من مواصفات تصميم معينة، فعلى سبيل المثال، قد يرغب أحد الناشرين الذين يقومون بإنشاء مقالات دورياتهم وفقاً لمعرف ما من معرفات أنواع الوثائق (DTDs)، أن ينقل هذه المقالات بطريقتين مختلفتين هما: عمل نسخة مطبوعة تتناسب مع متطلبات النشر التقليدي، وعمل صيغة أخرى تكون قابلة للعرض على الشاشة screen format للتوزيع عبر الإنترنت وللعرض على شاشات الحاسبات. وعلى ذلك فإن النسخة الأولى ستنقل في شكل يتم إرساله لآلات الطباعة، في حين تنقل الثانية في أحد الصيغ التي تدعمها متصفحات الويب، وهي عادة ما تكون لغة ترميز النصوص الفائقة (HTML) أو صيغة الوثيقة القابلة للنقل Portable Document Format (PDF).

وتحتاج هذه العملية أن تترجم التيجان البنائية في عملية الترميز إلى صيغ يمكن عرضها إما في شكل مطبوع أو على الشاشة، وتستخدم نماذج الكتابة Style sheets لهذا الغرض. فعلى سبيل المثال، قد يحدد معرف نوع الوثيقة عنصراً على أنه عنوان رئيسي من المستوى الثاني، ويشار إلى ذلك بالتاجين التاليين هكذا: <h2> </h2>. هذا في الوقت

الذي تشير فيه نماذج الكتابة إلى ذلك بالقول بأن هذا العنوان الرئيسي يجب أن يعرض بالخط المسمى تايمز رومان times roman على أن يكون حجم هذا الخط 13 نقطة، وبالخط (البنط) الثقيل، وأن يكون اتجاه النص من اليسار إلى اليمين. كما أنها يمكن أن تحدد بعض الخصائص الهامة الأخرى، مثل: المسافات الفاصلة بين الكلمات والسطور، وكيفية التعامل مع العناوين الرئيسية التي ترد قبل نهاية الصفحة. فنماذج الكتابة إذن تقدم تعليمات تفصيلية عن كيفية تحويل أي وثيقة يمكن تخيلها conceivable valid document والتي تم ترميزها وفقاً لأي من معرفات أنواع الوثائق DTDs.

وتمثل قوة اللغة المعيارية الموحدة (SGML) عقبة عند إنشاء نماذج الكتابة، فالقراء معتادون على الكتب جميلة التصميم، وتأتي الكثير من مظاهر هذا الجمال من مهارات الحرفيين الذين يقومون بعمليات بناء الصفحات وتصميمها ، ومعروف أن عملهم هذا ينطوي على قدر كبير من اللمسات الفنية، والعين البشرية حساسة بشكل كبير تجاه هذه اللمسات الفنية، مع أن تفاصيل كيفية إنشاء الجداول

الإحصائية وتنسيقها أو كيفية تصميم أحد الكتب الفنية وترقيمها لم تدون مطلقاً في شكل قواعد إجرائية أو آلية. وبسهولة يمكن لنماذج الكتابة أن تكون معقدة جداً، مع أنها لا تزال تحقق فشلاً في إرضاء عملية تحويل الوثائق المعقدة.

وحيث إن اللغة المعيارية الموحدة تمثل إطاراً عاماً فقد ظل الناس يعملون على إعداد نماذج الكتابة لأي نوع من معرفات أنواع الوثائق. وتسمى هذه النماذج "بلغة العناصر الدلالية لأسلوب الوثيقة و مواصفاتها Document Style Semantics and Specification Language (DSSSL) "، وبالطبع هي مهمة شاقة على أية حال، وحتى الآن، تتم كتابة برامج معالجة لغة العناصر الدلالية لأسلوب الوثيقة ومواصفاتها لبعض معرفات أنواع الوثائق البسيطة. ومع ذلك فالمهمة العامة تبدو طموحة جداً. وإنه لمن السهولة أن يتم إنشاء نموذج مرضٍ من نماذج الكتابة العلمية لمعرف واحد من معرفات أنواع الوثائق ويستخدم في سياق مفهوم فهماً جيداً. وهناك كتب ودوريات كثيرة تطبع عن طريق الترميز باستخدام اللغة المعيارية الموحدة مع نماذج الكتابة ذات الأغراض الخاصة ويمكن لنتائجها أن تكون مبشرة.

وفي سياق لغة ترميز النصوص الفائقة (HTML)، لم يكن هناك مفهوم رسمي لنماذج الكتابة ، حيث تقوم عملية

الترميز بدمج العناصر البنائية (مثل الأنواع المختلفة للقوائم) مع تيجان التنسيق (مثل تلك التي تحدد أحجام الحروف التي ترد بالأنباط الثقيلة أو بالأشكال المائلة). وتقدم التيجان توجيهات عامة بشأن المظهر، والتي تترجمها المتصفحات وتعمل على تكييف هذا المظهر حسب شاشة الحاسب عند الاستخدام، أما المظهر الذي يراه المستخدم فهو نتيجة عملية دمج وتوليف بين عملية الترميز التي يوفرها مصمم صفحة الويب، وضوابط التنسيق المبيتة داخلياً، والاختيارات التي يحددها المستخدم، وفضلاً عن ذلك يستطيع مؤلفو الوثائق بلغة ترميز النصوص الفائقة والراغبون في المزيد من التحكم في المظهر أن يدخلوا أنواعاً مختلفة من الخطوط ومن البرامج الصغيرة applets ومن أدوات الوصل الأخرى plug-ins.

وتبين اللوحة رقم (9-5) كلاًّ من نموذج عرض الصفحات المنسدل Cascading Style Sheets (CSS) ، ولغة الكتابة الموسعة Extensible Style Language (XSL)، اللذين يعدان من أساليب توفير نماذج عرض الصفحات في كل من لغة ترميز النصوص الفائقة، ولغة الترميز الموسعة (على هذا الترتيب). وعندما

أدرك مطورو لغة الترميز الموسعة (XML) أن أكبر تحد ستواجهه لغة الترميز الموسعة قبل أن تصبح مقبولة على نطاق واسع يتمثل في كيفية التحكم في مظهر الوثيقة، وعندما أدركوا ذلك، قدموا دعمهم لنموذج عرض الصفحة المنسدل (CSS)، ولغة الكتابة الموسعة (XSL)، ومن السابق لأوانه أن نعرف من سينجح، غير أن عملية الدمج بينهما قد تبدو مبشرة بالنجاح، حيث يتوافر لهاتين الطريقتين الإطار العام لوضع قواعد وأحكام دقيقة للإجراءات التي يجب اتخاذها عندما يتم تحديد نماذج عرض صفحات معينة، ولكن لا يوافق عليها المستفيد.

اللوحة رقم (9-5)

نموذج عرض الصفحات المنسدل، ولغة الكتابة الموسعة

إذا كانت لغات الترميز تصف العناصر البنائية للوثيقة، فإن نماذج عرض الصفحة تحدد كيفية ظهور هذه العناصر عندما تنقل لتعرض على الشاشة أو لتطبع على الورق. وقد تم تطوير نماذج عرض الصفحة المنسدلة للاستخدام مع إجراءات الترميز بلغة ترميز النصوص الفائقة، كما تعد لغة الكتابة الموسعة امتداداً لنماذج عرض الصفحات المستخدمة مع الترميز عن طريق لغة الترميز الموسعة. وإلى حد كبير تعد لغة الترميز الموسعة هذه إصدارة مبسطة من اللغة المعيارية الموحدة الكاملة التي توفر

إمكانية التحويل المبسط من لغة ترميز النصوص الفائقة، كما أن لغة الكتابة الموسعة هي مشتقة من لغة العناصر الدلالية لأسلوب الوثيقة ومواصفاتها، ويمكن لأي نموذج من نماذج الكتابة المصمم وفق نماذج عرض الصفحات المنسدلة أن يحول إلى لغة الكتابة الموسعة، وبطريقة آلية بحتة. وقد كان الأمل معقوداً في الأساس على أن تكون لغة الكتابة الموسعة مجموعة فرعية من لغة العناصر الدلالية لأسلوب الوثيقة ومواصفاتها، ولكن حدثت انحرافات عن هذا الهدف الأصلي، وما لغة الكتابة الموسعة الآن سوى مجرد مواصفة من المواصفات لا أكثر، وهناك أمل كبير في أن تحظى هذه اللغة بالقدر نفسه من الاهتمام الذي حظيت به لغة الترميز الموسعة.

وفي سياق نماذج عرض الصفحة المنسدلة، تحدد إحدى القواعد الأساليب التي سيتم تطبيقها على العناصر التي تم اختيارها من وثيقة معينة، مثال ذلك القاعدة البسيطة التي تقول : (Color: Blue)، فإنها تشير إلى أنه بالنسبة للعناصر المتوجة بالتاج h1 (وهو التاج المستخدم من قبل لغة ترميز

النصوص الفائقة للإشارة إلى العناوين الرئيسية من المستوى الأول) يجب أن تظهر ملونة باللون الأزرق، ولتوضيح ذلك بشكل رسمي، فإن كل قاعدة عادة ما تتكون من "مفتاح انتقاء selector" والذي يختار عناصر محددة في الوثيقة، وبيان declaration (وهو عادة ما يكون محصوراً بين قوسين هلالين، ويكون مكوناً من جزأين يفصل بينهما بعلامة الشارحة:colon ، حيث يشير أولهما للخاصية، في حين يشير الآخر للقيمة، وهذا البيان يشير إلى الأسلوب التحريري الذي يجب تطبيقه على العناصر.

أما نموذج الكتابة لنماذج عرض الصفحة المنسدلة فهي عبارة عن قائمة بالقواعد، وهناك العديد من الضوابط المختلفة التي تهدف إلى تبسيط أسلوب كتابة تلك القواعد، فعلى سبيل المثال، تشير القاعدة التالية التي ترد هكذا:

h1, h2 (font-family: sans-serif; color: blue)

إلى أن العنوانين الرئيسين الأول والثاني يجب أن يعرضا بلون أزرق، وبالخط المسمى sans-serif.

كما تحدد عملية الترميز لوثيقة مرمزة بلغة ترميز النصوص الفائقة البنية التي يمكن عرضها في شكل هرمي، ومن ثم فإن العناوين الرئيسية والفقرات والقوائم تعد جميعها من عناصر جسم body لغة ترميز النصوص الفائقة؛

بمعنى أن المداخل المدرجة بالقائمة تعد عناصر داخل القوائم، كما أن القوائم يمكن احتواؤها في قوائم أخرى، يضاف إلى ذلك أن القواعد في نماذج الكتابة لنماذج عرض الصفحة المنسدلة غالباً ما تتبع أو تأخذ أساليب من قواعد أخرى، فإذا لم تقم قاعدة معينة باختيار أحد العناصر العليا من الشكل الهرمي، مثال ذلك القاعدتان التاليتان المكتوبتان هكذا:

Body (font-family: serif)

H1,h2 (font-family: sans-serif)

فمن الواضح أن العنوانين الأول والثاني هما من عناصر جسم لغة ترميز النصوص الفائقة، ولكن لهما قاعدة صريحة تقول إنهما يجب أن يظهرا بالخط المسمى sans-serif، وحيث إنه لا يوجد في هذا المثال قاعدة صريحة خاصة بالفقرات أو القوائم، فإنهما يأخذان حكم الأساليب التي تطبق على الجسم الذي هو في المستوى الأعلى من الشكل الهرمي، وبناء على ذلك سيتم عرضهما بالخط المسمى serif.

إن نماذج عرض الصفحات يجب أن ترد مرتبطة مع

إحدى صفحات لغة ترميز النصوص الفائقة، ويتاح للمصمم في ذلك عدة خيارات منها تضمين النموذج في رأس الصفحة، أو عمل رابط link للملف الخارجي الذي يحتوي نموذج عرض الصفحات. وتجدر الإشارة إلى أن لكل متصفح نموذجه الضمني الخاص به لعرض الصفحات، والذي يمكن للمستفيد أن يعدلـه، كما يمكن أن يكون للمستفيد النموذج الخاص به.

ونظراً لأنه يمكن أن يكون هناك عدة نماذج لعرض الصفحات يمكن تطبيقها على الصفحة نفسها، فإن احتمالات التعارض قائمة طالما تعارضت القواعد، وقد تم تطوير آليات معتمدة على مبادئ بسيطة لتفادي هذه التعارضات، ومن أهم هذه المبادئ الأساسية التي تذكر في هذا الصدد أنه عندما تتعارض القواعد فإنه يتم اختيار واحدة وإهمال الأخرى، مع مراعاة أن تكون الأولوية لاختيار القواعد التي تحدد العناصر بصراحة، قبل تلك القواعد التابعة لقواعد أخرى inherited rules، وهناك مبدأ آخر يثير كثيرًا من الجدل، وهو أنه عندما تتعارض القاعدة التي يختارها المصمم مباشرة مع القاعدة التي يختارها المستفيد، تكون الأولوية للقاعدة التي اختارها المصمم، وعلى المستفيد الذي يرغب في تخطي هذه القاعدة وعدم الأخذ بها أن يرفع راية الأهمية أو مقولة: " هذا هو الأهم" important !" .

ومع أن هذا ليس بالحل الأمثل إلا أنه يسمح بتطوير نماذج عرض صفحات خاصة، مثل تلك الخاصة بتطوير صفحات تتفق واحتياجات ضعفاء البصر من المستفيدين الذين يرغبون في أن تعرض النصوص في حروف طباعة كبيرة.

لغات وصف الصفحات Page Description Languages :

بما أن المنشئين والقراء يحرصون على إضفاء أهمية كبرى لمظهر الوثائق، فمن المرغوب فيه أن تكون هناك أساليب تحدد مظهر الوثيقة مباشرة وبدون ترميز بنائي، وبالرغم من أن الأساليب المستخدمة في إنجاز هذا الهدف تتفاوت تفاوتاً كبيراً في تفاصيلها، فإنها تجتمع على هدف أساسي واحد، وهو نقل المواد النصية بالقدر نفسه من الدقة والجودة الطباعية التي تظهر بها أفضل الوثائق المطبوعة بالطرق التقليدية، وليس هذا أمراً سهلاً، فقليلة هي الأشياء التي يشعر الإنسان بمتعة في استخدامها بقدر المتعة في التعامل مع الكتب المطبوعة طباعة جيدة. ولا شك أن عمليات الطباعة وتصميم الصفحات وصناعة الورق وأساليب التجليد قد مرت بتطورات متلاحقة عبر السنين حتى وصلت إلى هذا المستوى التي هي عليه الآن من حيث

صلاحيتها وجودتها. وقد طورت الطرق المبكرة لتنسيق النصوص لأغراض إنتاج المواد المطبوعة، بيد أن العرض على شاشات الحاسبات قد أصبح له القدر نفسه من الأهمية، ويلقي هذا الجزء من هذا الفصل الضوء على ثلاث من لغات وصف الصفحات، وهي: لغة تي إكس TeX، ولغة بوست سكريبت Postscript، ولغة صيغة الوثيقة القابلة للنقل Portable Document Format "PDF". ومن الملاحظ أن لكل لغة من هذه اللغات الثلاث أهدافها المختلفة، والتي تسعى إلى تحقيقها بأساليب متباينة كذلك، غير أنها جميعاً تلتقي عند نقطة واحدة، وهي أنها جميعاً بمثابة أساليب عملية تؤدي وظيفتها بصورة جيدة في سياق نظم الإنتاج.

لغة تي إكس TeX :

هي أقدم اللغات الثلاث لوصف الصفحات التي سبقت الإشارة، وقد طور هذه اللغة دونالد كنوث Donald Knuth في عام 1980م تقريباً، ويهدف من وراء هذه اللغة إلى توفير أساليب طباعية عالية الجودة، مع التركيز بشكل خاص على طباعة المواد ذات الصبغة الرياضية، وفضلاً عن الرموز الخاصة الكثيرة في هذه اللغة، فإن الرمز الرياضي فيها يعتمد على تعبيرات معقدة لا يمكن تمثيلها في العادة بسلسلة واحدة متتابعة من الرموز. كما أن هذه اللغة تعمل على توفير قواعد لترميز المعادلات الرياضية لسلسلة من

رموز آسكي عند إدخالها وتخزينها ومعالجتها باستخدام الحاسب، مع توفير تيجان تشير إلى شكل عرضها على الشاشة.

ويلجأ معظم المستفيدين إلى استخدام واحدة من حزمتي لغة تي إكس، وهما: تي إكس العادية plain TeX، ولا تي إكس La TeX. وتسعى كل واحدة من هاتين الحزمتين إلى تعريف مجموعة تيجان التنسيق التي تغطي المواقف التي تظهر عادة مع عملية جمع الحروف، وهناك نظام يرتبط ارتباطاً وثيقاً مع لغة تي إكس لتصميم أحجام خطوط (أبناط) الطباعة يعرف بما وراء الخط أو الميتافونط metafont، وقد بذل كنوث جهداً كبيراً لإنتاج إصدارات من حجم الخط (البنط) القياسي الذي أعده لمجموعة من النظم الآلية.

وسوف تظل لغة تي إكس عملاً غير مسبوق لأغراض إعداد المواد ذات الصبغة الرياضية، ومن ثم يشيع استخدامها على نطاق واسع لدى المؤلفين في مجالات الرياضيات والفيزياء وما يرتبط بهما من مجالات أخرى ذات

صلة.

لغة بوست سكريبت Post Script :

كانت لغة بوست سكريبت أول منتجات نظم أدوبي Adobe التي نتجت عن جهود شركة زيروكس في عام 1984 م، وهي لغة برمجة تستخدم لإنتاج المواد البيانية أو الرسومية graphic المطبوعة، وبالرغم من أن هناك القليل جداً ممن كتبوا برامج بهذه اللغة فإن هناك كثيراً من الحاسبات لديها إجراءات طباعة تأخذ بموجبها النص أو الرسوم البيانية وتستحدث لها برنامجاً معادلاً للغة بوست سكريبت، ويمكن لهذا البرنامج أن يُرسل إلى جهاز التحكم في الطابعة الذي يقوم بدوره بتنفيذ برنامج بوست سكريبت وينشئ سلاسل متتابعة لتشغيل الطابعة.

ولعل إحدى نقاط القوة التي تتميز بها لغة بوست سكريبت هو دعمها الصريح لأحجام الخطوط (الأبناط)، وقد سجلت الكثير من النجاحات المبكرة التي تحققت لحاسبات آبل ماكينتوش بفضل دمج تصميمات الصور الخرائطية bit-mapped على الشاشة مع طابعات بوست سكريبت التي وفرت قدراً كبيراً من جودة المخرجات والتي لم تكن موجودة من قبل إلا للحاسبات مرتفعة الثمن. ومع تقنيات الطباعة بالليزر وتقنيات العرض على الشاشة ثم تشكيل

الحروف من مجموعة نقاط صغيرة، حيث تستخدم طابعات الليزر البسيطة معدل 300 نقطة في البوصة الواحدة. كما يمكن لآلات جمع الحروف الطباعية أن تتمتع بمعدل يصل إلى ما يقرب من 1200 نقطة في البوصة أو يزيد، ولكن في معظم شاشات الحاسبات يكون معدل النقط التي تؤلف الصورة حوالي 75 نقطة في البوصة، وبالتالي فإن أحجام الخطوط التي تبدو جذابة على شاشات الحاسبات ليست كتلك التي تستخدم مع الطباعة، كما أن معدلات عرض النص على الشاشة يجب أن تكون ملائمة لمعدل النقاء الخشن أو الواسع coarse resolution، وفي العادة تختلف وظائف نظم التشغيل التي تعرض النص على الشاشة عن أوامر لغة بوست سكريبت المستخدمة في الطباعة، وما لم يبذل قدر كبير من الحيطة والحذر فإن ذلك قد يؤدي إلى اختلاف في سطور خالية، وغير ذلك من اختلافات غير مقبولة.

وعلى الرغم من أن لغة بوست سكريبت هي أساساً لغة مصممة لإخراج الرسوم البيانية، وعلى الرغم كذلك من أن أثرها الأول ارتبط بمجال عرض مخرجات الحاسب على

طبعات الليزر- فإن برامج بوست سكريبت استخدمت كذلك مع تطبيقات أخرى، ومن أحد استخدامات هذه اللغة اختزان وتبادل عروض أي نص أو مخرج بياني، ومع ذلك يمكن القول بأنها ليست هي اللغة المثالية لهذا الغرض لأن بها الكثير من الاختلافات، ولأن البرامج تحتوي كذلك على افتراضات عن قدرات الحاسب الذي سيقوم بتنفيذها.

لغة صيغة الوثيقة القابلة للنقل (بي دي إف PDF) :

استفادت شركة أدوبي من تجربتها مع لغة بوست سكريبت في إنشاء صيغة الوثيقة المحمولة أو القابلة للنقل، وهي صيغة أفضل لتخزين صور الصفحات في شكل قابل للنقل ويعمل مستقلاً تماماً عن أي حاسب آلي محدد، وتتكون صيغة الوثيقة المحمولة - التي تعد من أكثر لغات وصف الصفحات شهرة اليوم - من صيغة قوية، ومجموعة من الأدوات المصممة لإنشاء الوثائق وتخزينها وعرضها.

وبالرغم من شيوع استخدام صيغة الوثيقة المحمولة في نظم إدارة الوثائق التجارية، فإن بعض المكتبات الرقمية لا تزال مترددة في استخلاصها، ويرجع أحد أسباب هذا التردد إلى أسباب فنية. وهذه الوثيقة من أفضل ما يلائم تمثيل الوثائق التي نشأت في أصول حاسوبية computer originals،

كما تستطيع هذه الصيغة أن تختزن الصور المرسومة خرائطياً، وعلى الرغم من توفير شركة أدوبي لبرمجية التعرف إلى الحروف بصرياً (OCR) لأغراض إنشاء ملفات هذه الصيغة، فإن الكثير من خصائصها قد تفتقد عندما تستخدم لأغراض تخزين ملفات الصور، كما قد تكون هناك ملفات كبيرة الحجم على نحو غير مرغوب فيه، لذلك تفقُد الكثير من جوانب المرونة المطلوبة للمكتبات الرقمية.

وبالإضافة إلى ما سبق، ترفض بعض المكتبات الرقمية والأرشيفات الملفات المكتوبة بهذه الصيغة على اعتبار أن هذه الصيغة مسجلة لحساب الجهة التي أنتجتها [وليست مشاعاً]، وهناك توجس من أن اتخاذ قرار استخدام صيغ الوثيقة المحمولة في المكتبات الرقمية قد يكون أكثر عرضة لمخاطر المستقبل من استخدامها لصيغة تباركها إحدى جهات التقييس المعروفة، وهو على أية حال تبرير مضلل، حيث حققت صيغ الوثائق القابلة للنقل أول نجاحاتها في شركة أمريكا AMERICA، التي ترحب بالمنتجات التجارية المدعومة دعماً جيداً, وينبغي على المجتمعات الأكاديمية

والعلمية أن تدرك أن الصيغة التي ترعاها بالصيانة شركة ما قد يكون أكثر استقراراً على المدى البعيد من المواصفات القياسية الرسمية غير المدعومة بمنتجات جيدة أو قاعدة عريضة من المستفيدين.

وأخيراً تجدر الإشارة إلى أن معرف صيغة الوثائق القابلة للنقل قد نشر على نطاق واسع، كما أن الاستخدام واسع النطاق لهذه اللغة في التطبيقات التجارية يضمن توافر البرامج لها، حتى ولو توقفت شركة أدوبي عن عملها تماماً أو توقفت عن دعم هذه البرامج.

اللوحة رقم (9-6)

لغة صيغة الوثيقة القابلة للنقل **Portable Document Format (PDF)**

صيغة الوثيقة المحمولة أو الوثيقة القابلة للنقل لشركة أدوبي من الصيغ الهامة، كما أنها صيغة مثيرة للاهتمام على اعتبار أنها نموذج لقدرة إحدى الشركات على استحداث إحدى المعايير أو إحدى المواصفات القياسية، والعمل على إتاحة هذه المواصفة على المستوى الدولي، وبالرغم من ذلك، لا تزال هذه الشركة تحقق أرباحاً جيدة. وصيغة الوثيقة القابلة للنقل هي عبارة عن إحدى صيغ تمثيل الوثائق، وهي ترد مستقلة عن برامج التطبيقات وعن الأنظمة الآلية. وتتكون الوثيقة في صيغة الوثيقة القابلة

للنقل من عدد من الصفحات المؤلفة من نص ورسوم بيانية وصور وغيرها من المعلومات المصاحبة. ومع ذلك فإن الخصائص التي ترتبط بصيغة الوثيقة القابلة للنقل يمكن أن تتسع لما أهو أكثر من مجرد النظرة الطباعية الجامدة للصفحة، حيث تدعم هذه الصيغة الروابط الفائقة، وأساليب البحث، وغيرهما من الخصائص الأخرى التي لا يمكن توافرها إلا إلكترونياً.

وصيغة الوثيقة القابلة للنقل هي تطور للغة البرمجة بوست سكريبت (التي طورتها شركة أدوبي كذلك)، وتتمثل إحدى طرق إنشاء ملف بصيغة الوثيقة القابلة للنقل في تغيير مسار divert سيل المعلومات التي عادة ما تأخذ طريقها إلى الطابعة. وعلى نحو بديل يمكن تغيير الملف من لغة بوست سكريبت أو من أي صيغة أخرى، ثم يتم بعد ذلك اختزانه وبثه عبر شبكة ما، أو عرضه على الشاشة، أو طباعته على الورق.

وتتمتع صيغة الوثيقة القابلة للنقل بكثير من نقاط القوة الفنية، حيث تبدو معظم الوثائق بهذه الصيغة عند عرضها

على شاشة الحاسبات شديدة الوضوح، ومع ذلك فبقدر تصميمها الجيد تكون جودة مظهرها عند الطباعة. وفيما عدا الملفات التي تحتوي على الصور المرسومة خرائطياً فإن الملفات في هذه الصيغة عادة ما تكون متوسطة الحجم، وإذا لم يكن بالحاسب الذي تعرض عليه الوثيقة المكتوبة بهذه الصيغة أنواع الأبناط الطباعية التي استخدمت في إنشائها فإن واصفات (الأبناط) font descriptors تمكن عارض الوثائق بصيغة الوثيقة القابلة للنقل من إنشاء (بنط) اصطناعي لها يكون عادة قريب الشبه من (البنط) الأصلي.

وعلى الرغم من أن هذه اللغة تعاني من صعوبة التمييز بين أنواع معينة من الملفات والعمل بأنواع غير معتادة من الأبناط، إلا أن صيغة الوثيقة القابلة للنقل باتت هامة لعرض الوثائق المتاحة على الإنترنت أو الوثائق المباشرة online documents وتوفر هذه اللغة كذلك العديد من الخصائص التي يطمح إليها المستفيدون، منها إمكانية إنشاء روابط فائقة (سواء روابط لعناصر داخل الوثيقة أو روابط تحيل إلى محددات موحدة خارجية لمواقع المصادر External URL)، يضاف إلى ذلك أن عارض الوثائق بصيغة الوثيقة القابلة للنقل يوفر أداة للبحث عن الكلمات داخل النص، مع أن البحث داخل الوثائق يمثل مشكلة، ومن خصائصها كذلك

دعمها للشروح ودعمها الجيد كذلك للطابعات، وكذلك توجد طريقة بهذه اللغة يستطيع منشئو الوثيقة من خلالها أن يمنع المستفيدين من طباعتها، أو من استخدامها بطرق ليست معتمدة من قبله.

وبعد فترة من التردد في تقديمها، قامت شركة أدوبي بتسويقها على استحياء، وقد تم ذلك مرحلياً عن طريق عمل برامج عرض جيدة لملفات هذه اللغة لجميع أنواع الحاسبات، وإتاحتها مجاناً على الإنترنت، وتحقق هذه الشركة أرباحاً ممن يقومون بشراء المنتجات التي تستخدم في إنشاء ملفات في صيغة الوثيقة القابلة للنقل، وليس ممن يستخدمون هذه الملفات. وهكذا يتضح لنا كيف أن شركة كهذه تستطيع أن تستحدث معياراً أو مواصفة قياسية وتتيحها للعالم، ومع ذلك لا تزال تحقق أرباحاً جيدة. ومع أن شركة أدوبي تمتلك حقوق الملكية الخاصة بمواصفة لغة صيغة الوثيقة القابلة للنقل، فإنها ومن أجل زيادة الدعاية لاستخدامها كشكل من أشكال تبادل المعلومات، تعطي تصريحاً غير مشروط لاستحداث ملفات

بهذه الصيغة، ولكتابة تطبيقات تنتج أعمالاً مكتوبة بهذه الطريقة، ولكتابة برمجية تقرأ الملفات المكتوبة بها وتجهزها. وإذا كان منافسوها قد تمكنوا من استحداث منتجات تقل أسعارها عن أسعار منتجات الشركة فإنها قد قبلت تلك المخاطرة.

بنية الوثائق في مقابل مظهرها :

استهل هذا الفصل بمناقشة متطلبين من متطلبات اختزان الوثائق في المكتبات الرقمية وهما: تمثيل بنية الوثائق، وتمثيل مظهرها. ولا ينبغي النظر إلى هذين المتطلبين على أنهما بديلان بعضهما عن بعض، أو على أنهما منافسان بعضهما لبعض، ولكن ينبغي النظر على أنهما مكملان أو متلازمان، ويستحق كلاهما القدر نفسه من الرعاية والاهتمام، وبالرغم من أن تمثيل أحدهما يلزم معه تمثيل الآخر في بعض التطبيقات، فإن عددًا من المكتبات الرقمية تحرص على اختزان إصدارتين من كل وثيقة. وإذا كانت الوثائق النصية تحتل مركز القلب من المكتبات الرقمية ومن صناعة النشر الإلكتروني، وإذا كان المؤلفون والقراء شديدي الإلحاح على طلبهم لهذه الوثائق، فمن الممكن القول بأن الأساليب المناسبة لتلبية تلك الاحتياجات قد توافرت.

الفصل التاسع

الفصل العاشر
استرجاع المعلومات وما وراء البيانات الوصفية

إن تقديم العون للمستفيدين من أجل الوصول إلى المعلومات من الخدمات الجوهرية التي تقوم بها المكتبات الرقمية. ويبدأ هذا الفصل - وهو أحد فصلين عن هذا الموضوع - بمناقشة "ما وراء البيانات الوصفية descriptive metadata "، أي الفهارس والكشافات والمعلومات المركزة الأخرى التي تستخدم لوصف الكائنات بالمكتبات الرقمية، ثم يلي ذلك مناقشة للطرق التي يتم بها البحث في متون النصوص بهدف الوصول إلى معلومات محددة. ثم يتوسع الفصل التالي (الحادي عشر) في عرض المفاهيم التي نوقشت في سياق هذا الفصل (العاشر) لمناقشة عملية البحث اللامركزي ، أي البحث عن المعلومات المبثوثة عبر المجموعات المستقلة بعضها عن بعض أو المتناثرة في كثير من النظم الآلية.

ويركز هذان الفصلان على السبل التي يسلكها المستفيدون في البحث عن معلومات محددة، وإذا كان البحث المباشر يعد أحد الطرق أو الإستراتيجيات التي يستخدمها المستفيدون لاستكشاف المعلومات، فإن التصفح

- وهو الاستكشاف غير المنظم لمعلومة معينة - قد أصبح طريقة شائعة وفاعلة لاستكشاف المعلومات غير المتوقعة. وإذا كانت المكتبات التقليدية تقوم بترتيب مجموعاتها ترتيباً موضوعياً في سبيل تيسيرها عملية التصفح، ويقدم تصنيف ديوي العشري وتصنيف مكتبة الكونجرس خطتين لتصنيف الموضوعات في بناء هرمي يمكن استخدامها في تنظيم مقتنيات المكتبات - فإن محرك البحث المعروف بياهو Yahoo - الذي يعد أوسع خدمات المعلومات انتشاراً على الإنترنت - يقدم خدمة تعد في أسسها خطة تصنيف للمواد المتاحة على الويب معززة بإمكانات البحث. وتتجه المكتبات الرقمية - بما يتوافر لها من الروابط الفائقة - لتطبيق الإستراتيجيات التي تجمع ما بين إمكانات البحث وإمكانات التصفح.

إن تنوع احتياجات المستفيدين من المكتبات الرقمية هو الذي يفسر لنا ما الذي يجعل عملية استكشاف المعلومات أمراً صعباً ومعقداً، كما يفسر لنا كذلك لماذا لا تعد طريقة واحدة من طرق استكشاف المعلومات ملائمة لإرضاء جميع

المستفيدين أو عدم ملاءمتها للبحث عن جميع المواد والموضوعات. وإن الهدف الأساسي من عملية "البحث الشامل comprehensive search" هو محاولة العثور على جميع المعلومات المرتبطة بموضوع معين، وعليه فإن الدراسات العلمية أو البحثية أو القانونية الجادة عادة ما تبدأ ببحث شامل للتعرف إلى الأعمال السابقة في مجالاتها. وفي حالة "البحث عن عنوان معروف a known-item search"، فإن المستفيد يبحث عن شيء محدد، مثل ذلك البحث عن الفصل الافتتاحي لكتاب موبي ديك Moby Dick، وغالبًا لا يكون المرجع الذي يبحث فيه على الوجه الأكمل، ومن ثم يتطلب الأمر بحثاً كاملاً للعثور على الشيء المطلوب، وبالرغم من ذلك، فثمة شيء محدد موجود، سوف يعرفه المستفيد عندما يعثر عليه. كذلك هناك ما يعرف "بالبحث عن الحقائق searches for facts "، فهذه الحقائق (مثل اسم عاصمة البربانوس Barbanos، أو الوزن الذري للزئبق) عبارة عن معلومات محددة متواجدة في كثير من مصادر المعلومات، ومن ثم فإن البحث عن الحقيقة لا يكتمل إلا بالعثور عليها. وبالرغم من أن مصدر الحقيقة يجب أن يكون جديراً بالتعويل عليه؛ فإن هناك كثيرًا من المصادر الأخرى المحتملة لوجود الحقائق. ومن هنا تبرز حاجة شائعة أخرى

هي الحاجة إلى "المعلومات العامة عن موضوع معين overview information on a topic"، وإن سؤالاً يقول: "كيف يستطيع المرء أن ينظم حفلة زفاف ؟ " أو "ما المقصود بالترميز بالمفتاح العام؟" وهذا قد يقود إلى كثير من المصادر المناسبة للمعلومات، ولكن المستفيد عادة ما يطمح إلى الوصول إلى شيء محدد، وبمجرد أن يعثر المرء على مادة هامة، فإنه يطمح إلى معرفة "المواد المرتبطة بها". وكثيراً ما يثور سؤال شائع أثناء إجراء أحد البحوث وهو: "هل هناك أي بحث لاحق يعتمد على العمل الذي تم طرحه في هذه المقالة ؟"

ما وراء البيانات الوصفية Descriptive Metadata :

إن كثيراً من طرق استكشاف المعلومات لا تقوم بالبحث الفعلي للكائنات في المجموعات، ولكنها تنطلق مما وراء البيانات الوصفية لتلك الكائنات. وما وراء البيانات الوصفية الخاصة بأحد الكائنات قد تكون مدخلاً من مداخل الفهارس، أو إحدى سجلات خدمات التكشيف، أو مستخلصاً من المستخلصات. وعادة ما يتم اختزان ما وراء البيانات

الوصفية هذه بصورة منفصلة عن الكائن نفسه التي تصفه، وإن كانت ترد في بعض الأحيان مضمّنة مع الكائن نفسه.

وعادة ما يعبر عما وراء البيانات الوصفية نصيًّا، مع أنها تستخدم لوصف كائنات صيغ غير نصية مثل الصور والتسجيلات الصوتية والخرائط وبرامج الحاسب. كما أن الفهرس الواحد قد يتضمن تسجيلات لجميع أنواع مصادر المعلومات وغيرها من الوسائط الأخرى والأشكال المختلفة، وهذا ما يكفل للمستفيدين من المكتبات الرقمية إمكانية استكشاف المواد في الوسائط كافة من خلال البحث في التسجيلات النصية.

وعادة ما يتم إنشاء سجلات ما وراء البيانات الوصفية عن طريق المهنيين، وتمثل فهارس المكتبات والكشافات العلمية استثمارات هائلة من عمل هؤلاء المهنيين المهرة، الذين تواصلت أعمالهم على مدى عقود أو حتى قرون من الزمن، ولهذه الحقيقة الاقتصادية أهميتها في فهم الاتجاهات الحالية؛ فمن ناحية أولى، تبدو أهمية البناء على الاستثمارات والخبرات الفنية القديمة، ومن ناحية أخرى، ثمة حافز قوي يدعو لإيجاد أرخص الأسعار وأسرعها كفاءةً لإنشاء ما وراء البيانات الوصفية، سواء تم ذلك عن طريق الفهرسة الآلية، أو عن طريق استخدام أدوات حاسوبية

تعمل على تعزيز الخبرة البشرية.

الفهارس :

تسجيلات الفهارس هي تسجيلات قصيرة تقدم معلومات مختصرة عن مقتنيات المكتبة، وتنسحب كلمة "فهرس" على التسجيلات التي تتمتع ببنية ثابتة وأن تكون منظمة وفق قواعد مقننة، أما المستخلص فهو عبارة عن ملخص قصير لوثيقة طويلة، أما الأشكال الأخرى من تسجيلات خدمات التكشيف فإنها لا تتمتع بالصفة الرسمية التي تكتسبها تسجيلات الفهارس، وإن كانت لها بنيتها المحددة أكثر من تلك التي تكون للمستخلص البسيط.

ومما تجدر الإشارة إليه أن فهارس المكتبات تؤدي وظائف كثيرة، ولا تقتصر وظيفتها على عملية استرجاع المعلومات وحدها؛ حيث يمكن لبعض الفهارس أن تقدم معلومات ببليوجرافية شاملة لا يمكن استخلاصها مباشرة من الكائنات، كما يمكن أن تتضمن معلومات عن المؤلفين، وعن الأعمال الفنية المتحفية، وفضلاً عن ذلك، ولأغراض تيسير التعامل مع المقتنيات تشتمل الفهارس على

معلومات إدارية، كالمعلومات التي تبين أماكن اختزان المواد (سواء أكان التخزين مباشراً على الإنترنت أم على أرفف المكتبة). وعادة ما تكون الفهارس أصغر حجماً من المقتنيات التي تمثلها، حيث يتم في المكتبات التقليدية وصف المواد التي تشغل عدداً كبيراً من الأرفف بسجلات يمكن أن توضع في عدد محدد من أدراج البطاقات الصغيرة التي توضع في مكان واحد أو في قاعدة بيانات مباشرة Online Database. كذلك يمكن عمل كشافات مطابقة Mirror لفهارس المكتبات الرقمية بهدف الارتقاء بالأداء وتوفير كثير من مقومات الثقة في تلك المكتبات.

وتنقسم المعلومات التي تتضمنها تسجيلات الفهارس إلى حقول رئيسية وحقول فرعية مع تيجان أو وسيمات تعرف بهذه الحقول، حيث يمكن أن يكون هناك حقل لاسم المؤلف مصحوباً بحقل فرعي باسم العائلة الخاص بهذا المؤلف، ويتم في هذا الصدد استخدام " القواعد الأنجلو الأمريكية للفهرسة"، و"صيغة مارك للفهرسة المقروءة آلياً "، لوصف العديد من مصادر المعلومات بأشكالها كافة، بما فيها المنفردات، والمسلسلات، والمواد الأرشيفية. وفي ظل العمل المكثف اللازم لإعداد تسجيلات فهارس تفصيلية، يتم الآن فهرسة المواد مرة واحدة (من قبل إحدى المكتبات

الوطنية كمكتبة الكونجرس)، ثم يتم بعد ذلك توزيع التسجيلات على المكتبات الأخرى عن طريق أحد المرافق المتخصصة، مثل مركز الحاسب الآلي للمكتبات على الخط المباشر المعروف بـ (أو سي إل سي) OCLC.

ويدور الآن جدل حول دور صيغة مارك للفهرسة المقروءة آلياً وقواعد الفهرسة المرتبطة بها في المكتبات الرقمية، ويتمثل هذا الجدل في التساؤل عن إمكانية الإفادة من الطرق التقليدية للفهرسة في دعم الصيغ الحديثة، وأنواع الوسائط الجديدة، وطرق النشر. وإذا كانت صيغة مارك لا تزال تحتفظ بأهميتها في سياق المواد التقليدية، فإن أمناء المكتبات قاموا بتوسعة مجال استخدامها للتعامل مع الأنواع الأحدث من الكائنات أو المواد التي تقتنيها المكتبات الرقمية، ومع ذلك فإن صيغة مارك هذه لم تتبنها أية منظمات أخرى غير المكتبات التقليدية.

خدمات التكشيف والاستخلاص :

يعتمد الباحثون في المجالات العلمية والتقنية على خدمات التكشيف والاستخلاص أكثر من اعتمادهم على

الفهارس، ويتوافر الآن لكل تخصص علمي تقريباً خدمة تكشيف واستخلاص تساعد الباحثين في هذا التخصص على الوصول إلى مقالات الدوريات العلمية، ومن أمثلة تلك الخدمات: قاعدة مدلاين Medline في الطب والأحياء، والمستخلصات الكيميائية Chemical Abstract في الكيمياء، وإنسبك Inspec في الفيزياء والحاسبات والمجالات ذات العلاقة بهما. وتقوم كل خدمة من هذه الخدمات بتكشيف المقالات المنشورة في أكبر مجموعة من الدوريات العلمية المتخصصة، ويشمل كل سجل خاص بأية مقالة على بيانات ببليوجرافية (كاسم المؤلف، وعنوان المقالة، وتاريخ نشرها، وغيرها من البيانات الهامة الأخرى)، هذا بالإضافة إلى معلومات أخرى تصف المحتوى الموضوعي لهذه المقالة والتي يتم تنظيمها بهدف تيسير عملية استرجاع المعلومات، وبرغم ما قد يكون من تفاوت بين الخدمات في تفصيل البيانات التي تقدمها فإن جميعها تتفق في قدر كبير من هذه البيانات. ولأن خدمات التكشيف والاستخلاص ظهرت في وقت كانت فيه الحاسبات بطيئة وغالية الثمن أكثر مما هي عليه الآن، فإن المعلومات التي تقدمها تلك الخدمات اعتمد في بنائها على دعم عمليات البحث النصي البسيط، ومع ذلك فقد أثبتت الأيام أهمية هذه التسجيلات وخاصة في كثير من النظم المرنة.

وغالباً ما يرغب المستفيدون في المجالات العلمية في الحصول على معلومات حول موضوع محدد، ولكن نظراً لاحتواء اللغة على كثير من الكلمات المتشابهة والمتداخلة والمترادفة في معانيها، فإن البحث الموضوعي لا يمكن التعويل عليه ما لم يكن هناك معلومات تكشيفية تصف المحتوى الموضوعي لجميع المواد. ويمكن للمعلومات التي تصف المحتوى الموضوعي أن ترد في شكل مستخلص، أو كلمات مفتاحية، أو رؤوس موضوعات، أو أية معلومات أخرى مناسبة. ومن الملاحظ أن بعض خدمات التكشيف والاستخلاص تطلب من المؤلفين أن يزودوها بكلمات مفتاحية أو مستخلصات للمقالات التي ينشرونها بالدوريات، إلا أن ذلك يؤدي في كثير من الحالات إلى عدم التناغم بين المصطلحات وعدم الثبات أو الاطراد في استخدامها، والأفضل من ذلك أن يتولى المكشفون المحترفون مهمة اختيار الكلمات المفتاحية أو الموضوعية الخاصة بكل مقالة من تلك المقالات.

وثمة طريقة فعالة غير أنها مكلفة، وهي استخدام أسلوب اللغات المقيدة Controlled Vocabulary، فحيثما تكون

هناك عدة مصطلحات تصف مفهوماً واحداً، يجب استخدام مصطلح واحد منها فقط وبشكل ثابت، ومن ثم يتم تزويد المكشفين بقوائم بالمصطلحات الموضوعية المقننة وقواعد استخدامها، وعدم استخدام أية مصطلحات أخرى من خارج تلك القوائم، وتلك هي الطريقة التي تتبعها مكتبة الكونجرس في قائمة رؤوس الموضوعات التي وضعتها، والمكتبة الوطنية الطبية في قائمة رؤوس الموضوعات الطبية التي تستخدمها والتي تعرف بـ " مش Mesh ".

ويتطلب تطبيق أسلوب اللغات المقيدة وجود مكشفين مدربين، كما يتطلب أيضاً مستخدمين مهرة؛ لأن المصطلحات التي يجب استخدامها عند صياغة استراتيجيات البحث في أي نظام استرجاع، يجب أن تكون متوافقة بل متطابقة مع المصطلحات التي قام المكشفون بتحديدها عند عملية التكشيف. وتجدر الإشارة إلى أن مجال الطب في الولايات المتحدة يتمتع بحظ وافر في هذا الصدد لما يتوافر له من كوادر متخصصة في مجال الخدمة المرجعية المتخصصة، وهناك اتجاه عام يسود في عالم المكتبات الرقمية يسعى إلى تزويد المستفيدين بالأدوات والأساليب التي تمكنهم من الوصول إلى المعلومات مباشرة وبدون مساعدة من اختصاصي المراجع، ويمكن الإفادة من تجربة قائمة رؤوس الموضوعات الطبية (مش) Mesh، ومكنز

الفنون والعمارة the arts and architecture thesaurus، في عملية ربط المصطلحات التي يقدمها المستفيدون عند التساؤلات بالمصطلحات التي تستخدم في مرحلة التكشيف.

اللوحة رقم (10-1)

قائمة رؤوس الموضوعات الطبية **Medical Subject Heading**

تقدم المكتبة الوطنية الطبية National Library of Medicine (NLM) خدمات استرجاع المعلومات في الطب و المجالات المرتبطة به منذ الستينات الميلادية، وحيث إنه في مجال الطب - وهو مجال معقد يعج بكثير من المصطلحات - يمكن أن يوصف المفهوم نفسه بمصطلحات علمية مختلفة أو بأية كلمات عامية دارجة مختلفة كذلك، فقد قامت المكتبة الوطنية الطبية بتطوير لغة مقيدة تعرف باسم قائمة رؤوس الموضوعات الطبية واختصاراً بمش "Mesh"، وهي توفر قائمة برؤوس الموضوعات التي تستخدم في تكشيف المقالات التي تقوم المكتبة بتكشيفها سنوياً والتي تصل إلى ما يقرب من 400.000 مقالة، وفي الفهرسة الموضوعية لجميع الكتب التي تقتنيها. كما تستخدم هذه

القائمة في عملية استرجاع المعلومات، وعلاوة على هذه القائمة - التي تضم ما يقرب من 18.000رأس موضوع - يوجد هناك مكنز منفصل به حوالي 80.000 مصطلح كيميائي. وقد رتبت المصطلحات في هذه القائمة في تسلسل هرمي، بحيث تأتي المصطلحات العامة، مثل "التشريح"، "و أعضاء جسم الإنسان"، و"الأمراض"، في قمة هذا الهرم، وفي المرتبة التالية من التسلسل الهرمي نجد مثلاً كل مصطلح يتم تقسيمه هو الآخر بشكل تسلسلي، فنجد مثلاً مصطلح التشريح يقسم إلى ستة عشر موضوعاً، تبدأ بمناطق الجسم، والهيكل العظمي والعضلي، وبعد ذلك يتم تقسيم مناطق جسم الإنسان إلى أجزائها المختلفة مثل البطن والأطراف والظهر وهكذا، ثم يتم بذلك تقسيم كل جزء إلى أجزائه الفرعية وهكذا. وقد زودت هذه القائمة بآلاف من الإحالات بمختلف أنواعها للربط بين المصطلحات.

إن نجاح قائمة رؤوس الموضوعات الطبية يعتمد على كفاءة المهنيين الذين يقومون على تحديثها وتطويرها، وعلى كفاءة أولئك الذين يقومون باختيار المصطلحات الموضوعية التي تعبر عن المحتوى العلمي للوثائق، كما يعتمد الاستخدام الناجح لهذه القائمة على كفاءة المستفيدين

واختصاصيي المراجع الذين لديهم الخبرة الكافية بمجال الطب ولديهم القدرة على صياغة استفساراتهم عن طريق المصطلحات والتراكيب اللغوية المناسبة المستخدمة في القائمة.

اللوحة رقم (10-2)

مكنز الفنون والعمارة The Art and Architecture Thesaurus

قامت شركة جي باول جيتي تراست J. Paul Getty Trust بإعداد وتطوير مكنز الفنون والعمارة كلغة مقيدة ليستخدم في وصف المعلومات واسترجاعها في مجالات الفنون الجميلة والعمارة، وفنون الديكور، والثقافة المادية. ويضم هذا المكنز ما يقرب من 120.000 مصطلح تعبر عن الكائنات والمواد النصية والصور وفنون العمارة والثقافة من جميع العصور ومن جميع الشعوب والثقافات، مع التركيز على الحضارة الغربية. ويمكن للأرشيفات والمتاحف والمكتبات أن تستخدم هذا المكنز لوصف مقتنياتها من تلك المواد التي يغطيها هذا المكنز، كما يمكن استخدامه عند البحث عن مثل هذه المواد.

وقد بدأ العمل الجاد في إعداد هذا المكنز في أوائل الثمانينات الميلادية، عندما كانت الإنترنت لا تزال في بدايتها، ومع ذلك فقد تم إنشاء المعلومات في صيغ مرنة تسمح باستحداث العديد من النسخ من المعلومات نفسها، كنسخة تتاح على الإنترنت، ونسخة تصدر مطبوعة، وصيغ حاسوبية أخرى متنوعة، وقد حرصت شركة جي تي تراست على تنظيم هذا المكنز في شكل ييسر استخدامه من قبل البرامج الآلية لأغراض استرجاع المعلومات ولمعالجة اللغة الطبيعية.

ويتم تقسيم هذا المكنز إلى سبع فئات، تشتمل كل فئة على سلاسل هرمية من المصطلحات، وهذه الفئات هي: المفاهيم المشتركة، والخصائص المادية، والأساليب والعصور، والوكلاء، والأنشطة، والمواد، والكائنات، ويتم تمثيل كل مفهوم بسلسلة متتابعة من المصطلحات، يعامل واحد منها على أنه "الواصف descriptor " (أي المصطلح المفضل). وتجدر الإشارة إلى أن هذا المكنز لا يكتفي بتقديم مصطلحات تعبر عن الكائنات أو المواد فقط، ولكنه يشتمل أيضاً على رؤوس الموضوعات أو الكلمات المفتاحية التي تستخدم في وصف هذه الكائنات مثل : أسلوب هذه الكائنات، والفترة الزمنية التي أنتجت فيها، والشكل الذي قدمت فيه، ولونها، وتشكيلها، واستخدامها.

كما يشتمل كذلك على مفاهيم علمية مثل النظريات والنقد.

لاشك إن تطوير وتحديث مكنز متخصص ضخم كهذا العمل يحتاج تكاليف باهظة، وخاصة أن المصطلحات في مثل هذه العلوم الناضجة تتغير باستمرار، ويضطر الاختصاصيون الفنيون في هذه المجالات إلى دعم التقنية الحديثة. ومن اللافت للانتباه، أنه على الرغم مما تنعم به شركة جي تي تراست من ثراء باذخ، فقد استغرق تطوير هذا المكنز سنوات طويلة.

دبلن كور Dublin Core :

منذ عام 1995م تقوم مجموعة دولية يقودها ستوارت وايبل Stuart Weibel من مركز الحاسب الآلي للمكتبات الخط المباشر OCLC بالعمل على استنباط مجموعة من عناصر ما وراء البيانات البسيطة التي يمكن تطبيقها على مجموعات كبيرة متنوعة من مواد المكتبات الرقمية، وتعرف مجموعة العناصر الأساسية التي اتفقت عليها تلك الجماعة "بدبلن كور Dublin Core"، وقد سميت بهذا الاسم نسبة إلى مدينة دبلن في ولاية أوهايو الأمريكية التي تم فيها أول اجتماع

لتلك الجماعة، وقد شارك عدة مئات من المتخصصين في ورش العمل التي أقيمت في دبلن، تبادلوا النقاش حول تصميم هذه المجموعة الأساسية من ما وراء البيانات عبر البريد الإلكتروني، ولعل روح التعاون التي خيمت على هؤلاء جميعاً تعتبر نموذجاً يحتذى به، ويبين كيف يمكن أن يعمل أفراد ذو اهتمامات مختلفة معاً. وتوضح اللوحة رقم (10-3) باختصار العناصر الخمسة عشر التي تم الاتفاق عليها لتمثل المجموعة الأساسية لما وراء البيانات.

ومما يلفت الانتباه أن تكون البساطة هي نقطة القوة في عناصر دبلن كور، وأن تكون في الوقت ذاته نقطة الضعف التي تعتريها؛ ففي الوقت الذي تتصف فيه القواعد التقليدية للفهرسة بالإسهاب والتعقيد وحاجة من يقوم بتطبيقها إلى برامج تدريبية ليتمكنوا من فهمها، نجد أن دبلن كور تتصف بالبساطة، وبالرغم من ذلك، تتناقض البساطة مع الدقة، وقد بذل فريق العمل جهداً كبيراً من أجل التوفيق بين هذين النقيضين. وتجدر الإشارة إلى أن الهدف الأساسي من العمل كان استحداث مجموعة واحدة من عناصر ما وراء البيانات يستخدمها الأشخاص غير المتدربين في وصف المواد التي يرغبون في نشرها إلكترونياً على الإنترنت. وبالرغم من أن هناك بعض الأشخاص لا يزالون متمسكين بوجهة النظر الدنيا هذه، فإنهم يتطلعون إلى رؤية

مجموعة بسيطة من القواعد التي يستطيع أي شخص أن يطبقها، وهناك آخرون يفضلون الثمرات التي تحققت من قواعد فهرسة أكثر تقييداً، وهم مستعدون لقبول المزيد من العمل الإضافي وتحمل التكاليف، ويستندون في وجهة نظرهم هذه إلى القول بأن البنية الإضافية للعناصر سوف تثمر عن المزيد من الدقة في تسجيلات ما وراء البيانات، فعلى سبيل المثال، إذا كانت هناك مداخل في أحد الحقول الموضوعية مستقاة من تصنيف ديوي العشري، فسوف يكون من المفيد تسجيل هذه المعلومة في تسجيلة ما وراء البيانات الخاصة بهذا المصدر، ولمزيد من الارتقاء بفاعلية ما وراء البيانات في عملية استرجاع المعلومات، يتوقع أن يكون للعديد من العناصر قوائم موصى بها من القيم، ومن ثم ستوجد مجموعة محددة من الأنواع، وعندها يمكن أن ينصح المكشفون بالاختيار من هذه القائمة.

وهناك استراتيجية مقترحة بتقديم خيارين هما: الخيار الأدنى أو الأساسي minimalist، والخيار البنيوي أو البنائي structuralist. بحيث يلتزم الخيار الأدنى بتحقيق المعايير

الأصلية التي تجعله قابلاً للاستخدام من قبل من لم يتلقوا تدريباً رسمياً على استخدامه. أما الخيار البنيوي فسيكون أكثر تعقيداً، ويحتاج في استخدامه إلى توجيهات كاملة وإلى موظفين مدربين.

اللوحة رقم (10- 3)

مجموعة دبلن كور **The Dublin Core**

تشكل العناصر الخمسة عشر التالية مجموعة ما وراء البيانات الأساسية لدبلن كور، ومن الملاحظ أن جميع هذه العناصر اختيارية كما يمكن تكرارها كلها، وقد تم اختصار وصف كل عنصر من هذه العناصر من تعريفات دبلن كور الرسمية بموجب تصريح من فريق التصميم) [على النحو التالي] :

العنوان : وهو الاسم الذي يعطيه المنشئ أو الناشر للمصدر.

المنشئ : هو الشخص أو الهيئة المسؤولة مسئولية كاملة عن المحتوى الفكري للمصدر، وهم المؤلفون في حالة مصادر المعلومات المكتوبة، والفنانون أو المصورون أو الرسامون في حالة المصادر البصرية.

الموضوع : هو موضوع المصدر، عادة ما يتم تمثيل

الموضوع بكلمة مفتاحية واحدة، أو بعبارة تصف المحتوى الموضوعي. وينصح باستخدام الكلمات المقيدة أو خطط التصنيف المقننة.

الوصف : هو وصف يرد في شكل نصي لمحتوى المصدر، بما في ذلك المستخلصات في حالة المواد شبه الوثائقية، وأوصاف المحتوى في حالة المصادر البصرية.

الناشر : هو الجهة المسؤولة عن إتاحة المصدر في صورته الحالية، مثل إحدى دور النشر، أو أحد الأقسام الأكاديمية بإحدى الجامعات، أو إحدى الشركات.

المشارك : هو شخص أو هيئة لم يرد ذكره في عنصر المنشئ، مع أنه قدم مساهمة فكرية جوهرية في المصدر، ولكن هذه المساهمة تعد ثانوية إذا ما قورنت بمساهمة المنشئ الأصلي، ومن أمثلة المشاركين: المحرر والمترجم والرسام.

التاريخ : هو تاريخ إنشاء المصدر أو إتاحته.

النوع : هو نوع المصدر، مثل : صفحة إنترنت، أو رواية، أو قصيدة، أو ورقة عمل، أو طبعة مبدئية من مقالة، أو

تقرير فني، أو مقالة مطولة، أو معجم.

الشكل أو الصيغة: هي صيغة بيانات المصدر أو شكلها، التي تستخدم لتحديد البرنامج وربما الأجهزة اللازمة لعرض المصدر أو تشغيله.

المعرِّف : هو عبارة أو كلمة [سلسلة من الحروف] أو رقم يستخدم لتحديد هوية المصدر تحديداً فريداً يميزه عن غيره، ومن أمثلة المعرِّفات الموجودة في الشبكات: المحددات الموحدة لمواقع المصادر URLs، والأسماء الموحدة للمصادر URNs.

المصدر : معلومات عن مصدر آخر اشتُقَّ منه المصدر الحالي.

اللغة : هي لغة المحتوى الفكري للمصدر الحالي.

العلاقة: هي معرف بمصدر آخر، ثم تحديد علاقته بالمصدر الحالي، ويسمح هذا العنصر بالربط بين المصادر القريبة أو المتصلة بعضها ببعض، وأوصاف المصدر التي سيتم ذكرها، وهناك مثالان على ذلك هما: إحدى طبعات عمل معين، وأحد فصول كتاب معين.

التغطية: هي وصف لخصائص التغطية المكانية والزمنية للمصدر.

الحقوق : بيان بإدارة الحقوق، أو عنصر تعريف مرتبط

بذلك البيان، أو عنصر تعريف مرتبط بخدمة تقدم معلومات عن إدارة حقوق المصدر.

التكشيف الآلي :

إن عمليتي الفهرسة والتكشيف مكلفتان عندما يقوم بهما مهنيون مهرة، وقد أثبتت التجارب أن كل سجل من سجلات الفهرسة قد تصل تكلفة إعداده وتوزيعه إلى ما يقرب من خمسين دولارًا، ومن الملاحظ أن الحاجة إلى المعلومات في بعض المجالات - كالطب والكيمياء - تكون متزايدة بدرجة كبيرة تبرر معها تلك النفقات الباهظة التي تصرف على عملية التكشيف الشاملة، وإن كان الوضع العام لا يقاس على هذين المجالين. وقد جرت العادة على أن تقتصر عملية فهرسة أحد الكتب المتخصصة على إنشاء بطاقة شاملة لهذا الكتاب دون إعداد بطاقات تفصيلية لكل موضوع من موضوعاته، كما أن معظم المواد الخاصة في المتاحف ودور المحفوظات والمكتبات لا تتم فهرستها أو تكشيفها على نحو مستقل.

وهناك العديد من المواد التي تحرص المكتبات الرقمية

على اقتنائها، غير أنه ليس هناك مبرر قوي لفهرسة كل مادة منها على نحو مستقل. فقد تكون المجموعات كبيرة جداً، كما تقف التغيرات المستمرة في طبيعة الكائنات الرقمية التي تشكل هذه المجموعات حائلاً دون تحمل استثمارات بعيدة المدى في فهرستها؛ فكل كائن رقمي قد تتعدد أشكاله ونسخه على نحو مستمر، بل إن الكائن الواحد قد يتألف من كائنات متعددة؛ وكل واحد منها يتغير على نحو مستقل، كما أنه في الوقت الذي تظهر فيه كائنات جديدة بشكل مستمر، تختفي فيه كائنات أخرى، وأحياناً ما يكون تصور المستفيد لأحد الكائنات على أنه نتيجة لتطبيق أحد البرامج الآلية، وإن ذلك قد يشهد تغيراً مع كل تفاعل يحدث مع هذا البرنامج. وهذه كلها عوامل تزيد من صعوبة فهرسة مواد المكتبات الرقمية وترفع من معدلات تكلفة هذه العملية.

ولهذه الأسباب، من المحتمل ألا تحظى عمليتا الفهرسة والتكشيف المهنيتان أو اللتان يقوم بهما مهنيون في المكتبات التقليدية بالقدر نفسه من الأهمية في المكتبات الرقمية. والبديل المطروح في هذا الصدد هو استخدام برامج آلية تقوم على إنشاء تسجيلات الفهرسة بشكل آلي، مع ملاحظة أن تلك التسجيلات التي يتم إنشاؤها آلياً

عادة ما تفتقر إلى مقومات الدقة الكاملة، وليس لها ميزة سوى أنها غير مكلفة، وسوف يضطر نظام البحث القوي إلى بذل جهد كبير من أجل تفادي تلك الرداءة التي تحيط بالتسجيلات الفردية، وقد أثبتت برامج بحث الويب التي تقوم بإنشاء كشافاتها بشكل آلي صدق هذه النتيجة، وإذا كان قد تبين أن سجلات الفهرسة ليست بهذا القدر من الجودة، فإن نجاح خدمات البحث أظهرت أن الكشافات هي المفيدة في هذا الصدد، فهي على الأقل تفضل البديل المتمثل في عدم وجود شيء على الإطلاق.

وتجدر الإشارة إلى أن كثيراً من آفاق التطور التي أدت إلى التكشيف الآلي جاءت نتيجة لجهود البحث العلمي حول عملية تصفح النصوص text skimming، حيث كانت الإشكالية متمثلة في كيفية تنظيم البريد الإلكتروني؛ فالمستفيد عادة ما يكون لديه كم هائل من رسائل البريد الإلكتروني. ويريد أن يعيد ترتيب هذه الرسائل وفقاً لموضوعاتها، ولكي يتم ذلك يتوقع أن يقوم البرنامج الآلي بتصفح هذه الرسائل وتحديد موضوع كل رسالة منها، وإذا عرفنا أن هذا الأمر

يصعب على البشر أن يقوموا به على نحو ثابت ومطرد، فإنه سيكون أصعب بالنسبة للبرامج الآلية، غير أن تقدماً مطرداً قد تحقق في هذا الشأن؛ حيث يمكن أن يقوم البرنامج بالبحث عن أدلة أو إشارات معينة clues في الوثيقة ـ والتي عادة ما تكون عناصر بنائية (مثل حقل الموضوع في رسالة البريد الإلكتروني) أو إشارات لغوية، أو كلمات مفتاحية بسيطة.

كذلك تعتمد عملية التكشيف الآلي على تلك الأدلة التي تتواجد في الوثيقة، ويتضح من ملاحظة اللوحة رقم (10-4) أن المثال الأول من الأمثلة التي تتضمنها، يعد نموذجاً ناجحا ً؛ لأن الوثيقة المتاحة على الإنترنت وفرت أدلة مفيدة، بحيث استطاع برنامج التكشيف الآلي في محرك البحث ألتافيستا AltaVista أن يحدد عنوان المقالة ومؤلفها، حيث اشتملت الصفحة - مثلاً - على العنصر التالي الذي ورد متوجاً على النحو التالي: <title> digital library concepts</title>

وقد قام المؤلف بإدراج هذه التيجان لتوجيه متصفح الويب عند عرض المقالة على الويب، كما أن هذه الطريقة كانت لها أهميتها كذلك حيث قدمت إرشادات لبرنامج التكشيف الآلي.

ومن الاستخدامات المحتملة للغات الترميز، مثل: اللغة

المعيارية الموحدة للترميز SGML، ولغة الترميز الموسعة XML، إمكانية استخدام التيجان البنائية في برامج التكشيف الآلي عند إنشاء تسجيلات تستخدم في أغراض استرجاع المعلومات؛ حيث يمكن أن تفسر منظومة الحروف التالية "ماري سيليست Marie celeste" عندما ترد في نص إحدى الوثائق على أنها إما اسم شخص، أو عنوان كتاب أو أغنية أو اسم سفينة، أو اسم ناشر، أو مسرحية، أو قد لا تكون اسماً على الإطلاق، أما عند استخدام التيجان البنائية، فإن منظومة الحروف يمكن أن تعرف وتوسم بما يدل على طبيعتها، وهكذا فإنه يمكن استخدام المعلومات التي تقدمها عملية الترميز في تمييز فئات محددة من المعلومات، مثل اسم المؤلف أو العنوان أو التاريخ.

إن عملية التكشيف الآلي تتسم بالسرعة وانخفاض التكلفة، ومع أن تحديد التكلفة الدقيقة يعد من الأسرار التجارية الناجحة، إلا أنه يمكن القول إن تكلفة إنشاء تسجيلة ببليوجرافية قد لا يتجاوز جزءاً من السنت، كما أن التكلفة التي تصرف على إعداد تسجيلة واحدة من قبل أحد

المفهرسين أو المكشفين المحترفين يمكن أن تساوي تكلفة إنشاء 100.000 تسجيلة أو أكثر يقوم بإنشائها أحد البرامج الآلية، كذلك يبدو مجدياً من الناحية الاقتصادية تكشيف أعداد ضخمة من المواد المتاحة على الإنترنت، بل حتى إعادة تكشيف هذه المواد على فترات زمنية متكررة.

ويمكن للقائمين على إعداد الفهارس والكشافات أن يوازنوا بين التكلفة والمزايا المحتملة، وتعد الطرق التقليدية التي تستخدم لأغراض إنشاء فهارس المكتبات أكثر صيغ ما وراء البيانات الوصفية ارتفاعًا في سعرها، أو من قبل خدمات التكشيف والاستخلاص. وفي هذا السياق يمكن القول بأن الخيار أو النموذج البنائي لمجموعة دبلن كور سيكون سعره معتدلاً مع احتفاظه بمعظم المزايا وتوفيره لبعض التكاليف، في حين سيكون الخيار أو النموذج الأدنى أو الأساسي من مجموعة دبلن كور أرخص سعراً لكنه غير متاح مجاناً، أما التكشيف الآلي فسيتمتع بالحدود الدنيا من درجات الدقة وبأقل التكلفة.

اللوحة رقم (10-4)

التكشيف الآلي **Automatic Indexing**

التسجيلتان التاليتان (اللتان تم إنشاؤهما عن طريق محـرك البحث المعـروف بألتافيسـتا في عـام 1997 إجـراء بعض التحريـرات الطفيفة عليهمـا) هما من تسجيلات التكشيف الآلي المنشأة آلياً بواسطة برامج البحث على الويب.

Digital library concepts

Key concept in the Architecture of digital library. William Y. Arms

…Corporation for National Research Initiatives Reston, Virginia

http;//www.dlib.org/dlib/july95/07arms.html-size 16K–7-Oc-95.
English

Repository References

Notice: HyperNews at union.ncsa.uiuc.edu. will be moving to a new machine and domain very soon. Expect interruptions. Repository
: References. This is a page
http;//union.ncsa.uiuc.edu/HyperNews/get/www./repo/refernce.html-size
5K- 12-May-95- English

يظهر النموذج الأول عملية التكشيف الآلي في أحسن

حالاتها؛ حيث يبين اسم مؤلف المقالة وعنوانها وتاريخ نشرها واسم المجلة الإلكترونية التي نشرت فيها. هذه البيانات - ولأغراض كثيرة - تعد بديلاً كافياً لإحدى التسجيلات التي يقوم بإعدادها أحد المكشفين المحترفين.

أما النموذج الثاني فيظهر بعض العيوب التي تكتنف عملية التكشيف الآلي، فمن الملاحظ أولاً أن من فهم محتوى الصفحة يرى أنها لا تستحق التكشيف، ويتضح من ناحية أخرى أنه إذا كانت المعلومات الواردة في هذه التسجيلة عن المكان والتاريخ صحيحة إلى حد ما، فإن العنوان يبدو غريباً، كما أن جسم التسجيلة ما هو - ببساطة - إلا الكلمات القليلة التي جاءت في بداية الصفحة.

إرفاق attachment ما وراء البيانات بالمحتوى :

يجب أن ترتبط ما وراء البيانات الوصفية بالمواد التي تصفها، وعادة كان في الماضي يتم اختزان ما وراء البيانات هذه على نحو منفصل، كفهرس أو كشاف خارجي، وبرغم مما كان لهذا الأسلوب من مزايا، إلا أنه كان يتطلب وجود روابط link بين ما وراء البيانات والكائنات التي تصفها، أما الآن فيلاحظ أن المكتبات الرقمية تتحرك نحو الاتجاه الآخر وهو اختزان ما وراء البيانات مع البيانات نفسها (سواء تم ذلك عن طريق دمج ما وراء البيانات مع الكائن نفسه، أو

إيجاد كائنين مرتبطين بعضهما ببعض ربطاً وثيقاً)، وهي على أية حال طريقة ملائمة بالنسبة للنظم اللامركزية، ولأغراض الحفظ الأرشيفي طويل الأمد؛ حيث إنها تضمن وصول البرنامج الآلي إلى كل من البيانات، وما وراء البيانات في آن واحد.

لقد خضعت آليات ربط ما وراء البيانات بصفحات الويب لكثير من الجدل، فبالنسبة للصفحة المرمزة بلغة ترميز النص الفائق HTML هناك طريقة مبسطة، وهي دمج ما وراء البيانات في الصفحة باستخدام التاج الخاص بصفحة النص الفائق وهو <meta> الذي يستمد من أي توصيف باستخدام لغة ترميز النص الفائق، لمجموعة عناصر دبلن كور الأساسية، ويبين الـجدول رقم (10-1) استخدام التيجان المشيرة لما وراء البيانات وهي <meta> (مع ملاحظة أن اختيار التيجان <meta> هو أحد قرارات تصميم النظام، كما أن مجموعة دبلن كور نفسها لا تحدد كيف يتم ربط ما وراء البيانات بالمادة التي تصفها.

الجدول رقم (10-1)

ما وراء البيانات المعروضة مع لغة ترميز النصوص الفائقة HTML

<meta> tags

Meta name = ”DC.subject” content=”Dublin core metadata element set”>

<”Meta name = ”DC.subject” content=”networked object description

Meta name = ”DC.publisher” content=”OCLC Online Computer Library <”.enter, Inc

<”. ,.Meta name=”DC.creator” content=”Weibel, Stuart L

weibel@oclc.org

Meta name=”DC.creator” content=”Miller, Eric J., .”> emiller@oclc.ort

Meta name=”DC.title” content=”Dublin Core Element Set Reference <”Page

<”Meta name=”DC.date” content =”1996-05-28

<”Meta name=”DC.form” scheme=”IMT” content=”text/html

<”Meta name=”DC.language” scheme=”ISO639” content=”en

< ”Meta name=”DC.identifier” scheme=”URL

Content =http://purl.oclc.org/metadata/dublin_core

وحيث إنه لا يمكن استخدام تيجان ما وراء البيانات meta tags مع ملفات أخرى غير الملفات المعدة بلغة ترميز النصوص الفائقة (HTML)، والتي أصبحت تمثل عائقاً على نحو متسارع، فإن عدداً من الهيئات العاملة مع اتحاد الويب WWWC قد طورت بنيةً أو هيكلاً أكثر عمومية يعرف بـ "إطار وصف المصدر the Resource Description Framework "RDF"، وهو ما تصفه اللوحة رقم (10- 5).

اللوحة رقم (10-5)

إطار وصف المصدر **The Resource Description Framework (RDF)**

إطار وصف المصدر هو إحدى طرق تبادل ما وراء البيانات ثم تطويره من قبل اتحاد الويب WWW Consortium بعد إفادته من المفاهيم التي أثمرت عنها جهود عديدة أخرى (منها صيغة بيكس PICS لتحديد مستويات البطاقات labels المحددة للخصائص: مثل العنف والصور الخليعة المتاحة على صفحات الويب)، وتضافر جهود كل من فريق دبلن كور ومصممي إطار وصف المصدر تضافراً كبيراً من أجل إنجاح هذا العمل.

ويمكن أن نعتبر خطة ما وراء البيانات - كما هو الحال مع دبلن كور- ذات ثلاثة جوانب هي: الجانب الدلالي semantics، والجانب البنائي syntax، والجانب التركيبي أو الإنشائي structure ؛ حيث يصف الجانب الدلالي كيفية تفسير المفاهيم، كما هو الحال مع مفهوم "التاريخ " و "المنشئ"، في حين يحدد الجانب البنائي كيفية التعبير عما وراء البيانات، أما الجانب

التركيبي فيحدد طبيعة العلاقات بين عناصر ما وراء البيانات (مثل: اليوم، الشهر، السنة، بصفتها مكونات لحقل التاريخ). ويقدم إطار وصف المصدر (RDF) نموذجاً بنائياً بسيطاً وعاماً لتحديد الجانب البنائي، كما أنه لا يشترط الأسلوب الدلالي للألفاظ المستخدم في خطة ما وراء البيانات، أما لغة الترميز الموسعة (XML) فتستخدم لوصف خطة ما وراء البيانات، ولأغراض تبادل المعلومات بين النظم الآلية وبين الخطط.

ويتكون النموذج البنائي من ثلاثة عناصر هي: المصادر Resources، وأنواع الملكية Property Type، والقيم Values. ولنأخذ مثلاً على ذلك العبارة البسيطة التي تقول: "إن شكسبير هو مؤلف مسرحية هاملت". ويمكن في خطة دبلن كور تمثيل ذلك على النحو التالي:

المصدر	نوع الملكية	القيمة
هاملت	منشئ	شكسبير
—	نوع	مسرحية

وتجدر الإشارة إلى أن يمكن لخطة أخرى لما وراء البيانات يمكن أن تستخدم كلمة "مؤلف author" بدلاً من

"منشئ creator"، وربما تستخدم كلمة "نوع type" بمعنى مختلف تماماً. ولذلك فإن الترميز اعتماداً على إطار وصف المصدر (RDF) قد ينص صراحة على أن ما وراء البيانات هذه يتم التعبير عنها بناءً لخطة دبلن كور هكذا:

```
<DC:creator>Shakespeare</DC:creator>

<DC:type>play</DC:type>
```

ولاستكمال هذا المثال، يجب تحديد كلمة "هاملت" بمزيد من الدقة، وإذا تمت الإشارة إلى ذلك عن طريق المحدد الموحد لموقع المصدر الخيالي imaginary URL الذي كتب هكذا: "http://hamlet.org"، فإن تسجيلة إطار وصف المصدر الكاملة والمستخدمة مع لغة الترميز الموسعة، ستكون على النحو التالي:

```
<RDF:RDF>
</RDF:description RDF:about = http://hamlet.org>
<DC:creator>Shakespeare</DC:creator>
<DC:type>play</DC:type>
<RDF:description/>
```

<RDF:RDF>

ومن الملاحظ أن الترميز في هذه التسجيلة يوضح بجلاء أن كلمتي "الوصف description"، و"عن about" قد تم تحديدهما في خطة إطار وصف المصدر، بينما تم تحديد "المنشئ" و "النوع" بأسلوب دبلن كور DC. وهناك خطوة إضافية واحدة مطلوبة لإكمال هذه التسجيلة، وهي: أن كلا من خطتي "إطار وصف المصدر"، و "دبلن كور"، يجب أن يعرفا على أنهما بمثابة [فراغات اسمية namespaces] في لغة الترميز الموسعة XML.

أما النموذج البنائي في إطار وصف المصدر فيسمح بأن تكون للمصادر أنواع ملكية تشير إلى مصادر أخرى، حيث يمكن على سبيل المثال أن تشتمل إحدى قواعد البيانات على تسجيلة عن شكسبير بها ما وراء البيانات عنه، مثل: متى عاش وأين ؟ وما الطرق المختلفة التي كان يتهجى بها اسمه ؟ أما نوع ملكية المنشئ في دبلن كور فقد يشير إلى هذه التسجيلة على النحو التالي :

</DC:creator RDF:about = http://people.net/WS>

وبهذه الطريقة يمكن بناء أوصاف ما وراء البيانات الصعبة والمعقدة جداً من مكونات بسيطة، فباستخدام إطار وصف المصدر الخاص بالجوانب الدلالية والبنائية،

وباستخدام لغة التهيئة الموسعة يمكن للنظم الآلية أن تقوم بربط ما وراء البيانات مع الكائنات الرقمية، كما تستطيع أن تقوم بتبادل ما وراء البيانات المأخوذة من خطط مختلفة.

أساليب استرجاع المعلومات :

مجال استرجاع المعلومات من المجالات التي شهدت تضافراً ولسنوات طويلة بين جهود كل من علماء الحاسبات ومهنيي المعلومات، كما أن هذا المجال لا يزال حتى الآن من الجبهات الخصبة لجهود البحث العلمي، هذا فضلاً عن أنه يعد واحدًا من مجالات المكتبات الرقمية القليلة التي توافر لها طرق البحث المنهجية systematic methods التي تستخدم في قياس معدلات أداء أساليب استرجاع المعلومات المختلفة.

مفاهيم أساسية ومصطلحات :

ثمة بعض المفاهيم البسيطة المتداولة في مجال استرجاع المعلومات، فهناك "الاستفسار "query"[1]، وهو عبارة عن سلسلة من الكلمات التي تصف المعلومات التي

[1] (1) يعرف في اصطلاح المكتبيين باستراتيجية البحث (المترجمان).

يبحث عنها المستفيد، وكل مصطلح يرد في هذه السلسلة يسمى "المصطلح البحثي search term"، والاستفسارات قد تتكون من مصطلح بحثي واحد، أو مجموعة من المصطلحات، التي يعبر عنها باللغة الطبيعية، أو بتعبير أسلوبي مستخدمًا رموز خاصة.

وعند الاسترجاع تقوم بعض أساليب استرجاع المعلومات بمقارنة [1] "كلمات الاستفسار" بكل كلمة يتضمنها النص بأكمله، وبدون إجراء أي عملية تحديد لوظيفة كلمات هذا النص، ويعرف هذا الأسلوب "ببحث النصوص الكاملة Full-text searching"؛ في حين تقوم بعض الأساليب الأخرى بتحديد مجموعة حقول ببليوجرافية أو عناصر بنائية (مثل حقل المؤلف، أو رأس الموضوع) وتقييد البحث في واحد أو أكثر من هذه الحقول (كأن يحدد البـحث هكذا author = Gibbon، أي البحث في حقل المؤلف عن الاسم Gibbon)، ويعرف هذا الأسلوب بالبحث الحقلي field searching [2]. ومن الملاحظ أن كلا الأسلوبين يملكان أسباب القوة في البحث، وغالباً ما تسعى أساليب البحث الحديثة في استرجاع

---

[1] (1) الأدق تسميتها بالمضاهاة matching وليس المقارنة (المترجمان).

[2] (2) يعرف بتسميات أخرى كالتقييد الحقلي، أو البحث المتقدم, وهناك أساليب متباينة تختلف من برمجية بحث لأخرى في تطبيق هذا الأسلوب (المترجمان).

المعلومات للجمع بينهما.

وبينما يحتاج أسلوب البحث الحقلي إلى طريقة ما لتحديد هذه الحقول، فإن أسلوب بحث النصوص الكاملة لا يحتاج إلى هذا المتطلب، وفي ظل ما يتوافر لأجهزة الحاسبات الحديثة من إمكانات قوية، يمكن لأسلوب بحث النصوص الكاملة أن يكون فعالاً، حتى في حالة بحث النصوص التي لم تسبق معالجتها، ومع ذلك تبقى النصوص غير المتجانسة ذات الأطوال والأساليب والمحتويات المتفاوتة صعبة في بحثها بالكفاءة الكافية، كما يمكن أن تأتي نتائج البحث غير مرضية.

وبالرغم من أن نظامي المعلومات القانونية "ويسلو Weslaw"، و"لكزس Lexis" بُنيا على أساس أسلوب بحث النصوص الكاملة، فإن معظم الخدمات -عند توافر ما وراء البيانات - تفضل استخدام أسلوب البحث الحقلي أو البحث الحر في المستخلصات وما وراء البيانات الأخرى.

وهناك بعض الكلمات[1] - مثل الضمائر، وحروف

[1] (1) تعرف بكلمات التوقف (المترجمان).

العطف، والأفعال المساعدة - التي تتكرر كثيراً في النص، ليست لها القيمة الكبرى التي تبرر استرجاعها، وعليه تسعى معظم النظم إلى عمل قائمة بالكلمات المعلقة التي يجب تجاهلها في الاستفسارات، وتسمى هذه القائمة "بقائمة التوقف والاستبعاد stop list". ويعد اختيار كلمات التوقف من المهام الصعبة، حيث يعتمد ذلك على لغة النص، وقد يرتبط ذلك بالمجال الموضوعي للنص، لهذا السبب وبدلاً من قائمة التوقف المحددة مسبقاً، تستخدم بعض النظم أساليب إحصائية لتحديد أكثر الكلمات عمومية أو شيوعاً من حيث استخدامها، وتجعل هذه الكلمات من الكلمات التي لا يتم استرجاعها. وحتى الآن ليس هناك نظام كامل، بل هناك دائماً خطر يبرز من احتمال رفض بعض الاستفسارات الهامة، مثل عبارة: "نكون أو لا نكون to be or not to be"، إذ يمكن لهذه الجملة عندما يراد البحث عنها أن ترفض، أو لا يتم استرجاع ما يتصل بها، لا لشيء إلا لأن جميع كلماتها من الكلمات التي حددت في قائمة التوقف.

ويعد استخدام ما يعرف بالملف المصنف أو المقلوب inverted file الأسلوب الآلي الأساسي المستخدم في مضاهاة مصطلحات البحث في مقابل مجموعة الوثائق النصية وهو ما توضحه اللوحة رقم (10-6).

اللوحة رقم (10-6)

الملفات المصنفة

الملف المصنف هو قائمة بالكلمات الموجودة في مجموعة من الوثائق وأماكن وجودها [1] في تلك الوثائق، وفيما يلي مثال لجزء صغير من ملف مصنف :

الكلمة	الوثيقة	الموقع
المعداد abacus	3	94
	19	7
	19	212
ممثل Actor	2	66
	19	200
	29	45
شجر الحور الهزاز aspen	5	43
الجزيرة المرجانية atoll	11	3
	34	40

ويتبين من هذا الجزء من الملف المصنف أن كلمة "المعداد" هي الكلمة رقم 94 في الوثيقة رقم 3، ورقم 7،

[1] (1) المفترض مرات تكرارها وليس أماكن وجودها لأن ذلك يتطلب أن تحدد أماكن وجود كل كلمة في النص (المترجمان).

ورقم 212 في الوثيقة رقم 19، وأن كلمة "ممثل" هي الكلمة رقم 66 في الوثيقة رقم 2، والكلمة رقم 200 في الوثيقة رقم 19، والكلمة رقم 45 في الوثيقة رقم 29، وهكذا.

كما يمكن استخدام الملف المصنف للتعرف إلى عدد مرات تكرار مصطلح بحثي معين في مجموعة من الوثائق. ففي المثال السابق سوف يبدأ البحث عن كلمة "ممثل" بالنظر في الملف المصنف، وسوف نجد تلك الكلمة في الوثائق التي تحمل الأرقام 2، 19، 29. وتعد الإشارة أو الرجوع البسيط للملف المصنف من العملية السريعة للحاسب الآلي.

وتسمى قائمة الأماكن الخاصة بكلمة معينة "بالقائمة المصنفة"، وتقوم معظم القوائم المصنفة هذه بتحديد مواقع الكلمة في الوثيقة، وهو أمر له أهميته عند عرض نتائج البحث، لاسيما في الوثائق الطويلة، بحيث يتم عرض الجزء المناسب من الوثيقة الذي وجدت فيه الكلمة التي يبحث عنها والتي تظهر مظللة.

وحيث إن الملفات المصنفة تحتوي على كل كلمة في مجموعة من الوثائق (ما عدا كلمات التوقف)، فإنها تكون ملفات كبيرة، فالملف المصنف لمواد مكتبة رقمية معينة قد

يقترب من نصف حجم جميع الوثائق الموجودة فيها حتى بعد عملية ضغط هذا الملف، ولذلك - وعلى حساب مساحة التخزين - يوفر الملف المصنف طريقة سريعة للتعرف إلى عدد مرات تكرار الكلمة الواحدة في مجموعة من الوثائق. ومن الملاحظ أن معظم طرق استرجاع المعلومات تعتمد على هذه الملفات المصنفة.

البحث البوليني أو المنطقي Boolean search :

يتكون الاستفسار المنطقي Boolean query من مصطلحي بحث أو أكثر مرتبطين فيما بينهما برابط منطقي Boolean operator مثل حرف: "و and "، و" أو or"، و "ماعداء أو بدون not". لاحظ الاستفسار التالي على سبيل المثال: Abacus andactor، الذي يمكن تمريره على الملف المصنف الموضح في اللوحة رقم (10-6)، مع ملاحظة أن مصطلحي البحث مرتبطان برابط منطقي. فعند تنفيـذ هذا الاستفسار يتم في المرحلة الأولى قراءة القوائم المصنفة بحثاً عن كلمة " abacus" (ويتبين أنها توجد في الوثيقة رقم 3، 19)، والبحث عن كلمة "actor" (والتي يتبين أنها توجد في الوثائق رقم 2،

19، 29)، أما المرحلة الثانية فتتمثل في المضاهاة بين القائمتين للتعرف إلى الوثائق المشتركة في كلتا القائمتين، إذ يتضح أن الكلمتين قد وردتا في الوثيقة رقم (19)، وبالتالي تكون هي الوثيقة الوحيدة التي تجيب عن الاستفسار.

ومما تجدر الإشارة إليه أنه عندما تكون القوائم المصنفة قصيرة تتم عملية البحث البوليني عن مصطلحات بحث قليلة بشكل سريع جداً، كما هو الحال عند بحث الاستفسارات البسيطة تقريباً، غير أن المتطلبات الحاسوبية تزداد بشكل كبير مع المجموعات الكبيرة من المعلومات ومع الاستفسارات المعقدة.

ويمكن أن تستخدم الملفات المصنفة لتوسعة مجال المفاهيم الأساسية لعملية البحث البوليني، ولأن مكان الكلمات داخل الوثائق يكون مسجلاً في القوائم المصنفة، فيمكن استخدامها للبحوث التي ترغب في التعرف إلى الموقع النسبي لكلمتين، مثل كلمة west بحيث تليها كلمة Virginia [1]. كما يمكن استخدامه أيضاً في عمليات "البتر truncation" أي البحث عن كلمات تبدأ بحروف معينة. في كثير من أنظمة البحث سيؤول البحث عن استفسار كُتب هكذا: comp?، إلى البحث عن جميع الكلمات التي تبدأ بهذه

[1] (1) يعرف هذا بالبحث التجاوري أو المكاني (المترجمان).

الحروف الأربعة C و O و m و p، ومن ثم يتم استرجاع جميع الكلمات التي تبدأ بهذا المقطع مثل: compute, computing
computer, computers, computation، ولكن ما يؤخذ على هذه الطريقة أنها لا تميز بين الكلمات المناسبة وغير المناسبة التي تبدأ بالحروف نفسها، حيث يمكن في هذا الاستفسار السابق استرجاع كلمة company.

الترتيب الطبقي لنتائج البحث ranking closeness of match :

يعد البحث البوليني من أساليب البحث القوية، لكنه يعتمد في الاسترجاع فقط على المضاهاة الكاملة، وبالتالي استرجاع الكلمات المتماثلة تماماً، فعند البحث عن كلمة Library (مكتبة) لا تسترجع كلمة Libraries (مكتبات)، كما أن الاسم John Smith لا يعامل على أنه J. Smith كاسم واحد، مع أن الجميع يقر بذلك. وهناك مجموعة كبيرة من الأساليب الفنية التي تحاول معالجة مثل تلك المشكلات.

وهناك أحد الأساليب الحديثة الذي يحاول عدم المضاهاة الكاملة بين الوثائق والاستفسار، ولكن بدلاً من ذلك يعتمد

على تحديد بعض مقاييس التشابه بين الاستفسار وكل وثيقة. فلنفترض مثلاً أن إجمالي عدد الكلمات المختلفة في مجموعة من الوثائق هو (n)، فيمكن تمثيل وثيقة معينة عن طريق متجه vector في الفراغ البعدي لعدد الكلمات n-dimensional space، وإذا كانت الوثيقة تحتوي على كلمة معينة فسيكون للمتجه القيمة (1) في البعد المناظر، وإلا سيكون له القيمة (صفر)، كما يمكن أيضاً تمثيل الاستفسار بمقدار قرب هذين المتهجين بعضهما من بعض، وهذا قد يقاس بقيمة الزاوية بين هذين المتجهين في الفراغ البعدي لعدد الكلمات، وبمجرد الانتهاء من حساب هذه المقاييس لكل وثيقة، يمكن ترتيب النتائج ترتيباً طبقياً أو ترتيباً تنازليًا من أفضل درجات المضاهاة إلى الأسوأ وهكذا. وهناك عدة أساليب خطية لهذا الترتيب تعتمد على أساس هذا المفهوم العام. وتستفيد عدة أنواع مختلفة من الطرق الاحتمالية من التوزيع الإحصائي للكلمات في المجموعة، وهذه الطرق تنبثق من ملاحظة أن الكلمات الدقيقة المختارة من قبل المؤلف لوصف الموضوع، أو من قبل المستفيد عند طرح الاستفسار، قد اختيرت من مجموعة من الاحتمالات، ولكن [يلاحظ] أن الكلمات الأخرى قد تكون ملائمة بالقدر نفسه.

معالجة اللغة الطبيعية واللغويات الحاسوبية :

ليست الكلمات في إحدى الوثائق مجرد سلاسل

عشوائية من الحروف، بل هي كلمات إحدى اللغات (كالإنجليزية مثلاً)، وهذه الكلمات تنظمها عبارات، وجمل، وفقرات. ومعالجة اللغة الطبيعية هي فرع من علوم الحاسب الذي يحاول استثمار إمكانات الحاسبات في تفسير الكلمات ومعالجتها على اعتبار أنها جزء من إحدى اللغات. وتعد برامج التصحيح الإملائي التي تستخدم مع برامج معالجة الكلمات من أشهر تطبيقات هذا الفرع؛ حيث تستخدم أساليب معالجة اللغة الطبيعية لاقتراح الكلمات البديلة للكلمات التي لا يتعرف إليها البرنامج.

وكان من إنجازات علم اللغة الحاسوبي - وهو فرع من علم اللغويات يتعامل مع قواعد اللغة واللغويات - تطوير برامج آلية تستطيع أن تُعرب أية جملة تقريباً وبدقة معقولة، ثم تضيف الكلمات حسب أنواعها (فعل، اسم، صفة... إلخ)، ثم تجميع الكلمات معاً في عبارات بسيطة ومركبة، كما تستطيع تحديد العناصر البنائية في الجمل (كالفاعل، والفعل، والمفعول به... إلخ). وفي سبيل تحقيق هذه الأهداف تطلب الأمر من علماء اللغة أن يجددوا فهمهم

لقواعد اللغة، وأن يعوا كثيرًا من الأمور الدقيقة التي قد لا توجد في كتب النحو التقليدية، ولقد أجريت أبحاث هامة في مجال استرجاع المعلومات مستخدمة الجمل الاسمية، حيث يمكن في كثير من النصوص أو السياقات التعرف إلى محتوى الجملة عن طريق استخراج الأسماء والجمل الاسمية من هذه السياقات وإجراء البحث فيها، مع ملاحظة أن هذا العمل لم يقتصر على اللغة الإنجليزية بل يتم تطبيقه على لغات أخرى كثيرة.

ومما تجدر الإشارة إليه أن عملية الإعراب تحتاج إلى معرفة أساسيات في الصرف، حتى يمكن التعرف إلى الصيغ المختلفة التي يمكن أن تأخذها الكلمة الواحدة، كما هو الحال في صيغ المفرد والجمع (ككلمة library and libraries) وتصاريف الأفعال (مثل الفعل look عندما يكون في صيغة المصدر والمضارع والماضي هكذا: Look, looks, looked)، ومن المفيد أحياناً بالنسبة لعملية استرجاع المعلومات الحد من الصيغ المختلفة للكلمة إلى الجذر العام لها، واستخدام هذا الجذر كمصطلح بحثي، وهو ما يعرف في مجال البحث "بالتجذير stemming"، وهو أسلوب في البحث أكفأ من أسلوب "البتر truncation"؛ لأن هذا الأسلوب - أي التجذير - يقوم بفصل الكلمات ذات المعاني المختلفة تماماً (كما هو الحال مع

كلمتي company و computer) [1]، هذا في الوقت الذي يقوم فيه بالتعريف بكلمتي computing و computer على أنهما صيغتان مختلفتان من الجذر نفسه. وفي اللغة الإنجليزية - وهي لغة دائماً ما يكون فيها جذر الكلمة في البداية تقريباً - يمكن إجراء البحث بالتجذير عن طريق بتر الكلمات، وربما عن طريق إجراء بعض التعديلات على الحروف القليلة الأخيرة من الكلمة، كما يكون أيضاً في بعض اللغات الأخرى - كالألمانية - من الضروري إجراء بعض التعديلات على بداية الكلمات.

لقد طور علماء لغويات الحاسبات مجموعة من المعاجم والأدوات الأخرى التي تستخدم لأغراض معالجة اللغة الطبيعية (منها المعاجم lexicon والمكانز thesauruses). والمعجم يشتمل على معلومات عن الكلمات والصيغ المختلفة لها، واستخداماتها النحوية، أما المكنز فيقوم بربط الكلمات على أساس معانيها، وبعض هذه الأدوات تكون

[1](1) هاتان الكلمتان يتم استرجاعهما عند استخدام البتر لكلمة (المترجمان).

لغايات عامة، وقد تكون لغايات متخصصة في مجال موضوعي محدد (كما هو الحال مع مكنز الفنون والعمارة، وقائمة رؤوس الموضوعات الطبية).

ويمكن للمعاجم والمكانز أن تزيد من كفاءة عملية استرجاع المعلومات من خلال تعريفها للكلمات بأنها شيء أكثر من مجرد سلاسل عشوائية من الحروف، حيث يمكنها - على سبيل المثال - أن تعرف المترادفات (مثل كلمتي "سيارة وعربة "car" and "automobile")، وأن تقوم بربط مصطلح عام بحالة معينة (مثل كلمتي "علوم" وعلم الكيمياء science and chemistry)، أو تقوم بربط مصطلح فني بما يعادله في اللغة العامية (مثل مصطلح "جمجمة" و "مخ" cranium and brain).

وبشكل عام يمكن القول بأن إعداد المعاجم والمكانز ليست بالأمر الهين بل من الأعمال الصعبة، ومهما أُتقِن العمل في إعدادها فلا توصف بالكمال إطلاقاً؛ لأن اللغة سريعة التغير، كما أن المصطلحات متجددة بشكل مطرد وخاصة في مجالات البحث الحيوية.

واجهات المستفيدين ونظم استرجاع المعلومات :

تعتمد نظم استرجاع المعلومات في إظهار كفاءتها على

مدى قدرة المستفيدين على استثمار كل إمكانات البحث التي توفرها؛ فعندما يكون المستفيد هو أمين مكتبة طبية مدرباً، أو محامياً تدرب على نظم البحث أثناء دراسته فإنه يمكن لهما استثمار تلك الإمكانات استثماراً كاملاً، والمستفيدون غير المدربين هم أقل كفاءة من نظرائهم المدربين سواء عند صياغة استفساراتهم أو عند تقديرهم لنتائج البحوث التي يجرونها.

ويتمتع كل من أسلوبي الاتجاه الفراغي والاحتمالات في عملية استرجاع المعلومات بكفاءة كبيرة عند التعامل مع الاستفسارات الطويلة، وهناك تجربة مثيرة تتمثل في استخدام استفسار طويل جداً، ربما لطول مستخلص إحدى الوثائق، وهو ما يعادل استفسارًا موجهًا للحاسب باسترجاع الوثائق التي تضاهي هذا المستخلص، وهناك كثير من نظم البحث الحديثة تميزت بكفاءة واضحة عندما تم اختبارها، مع أن الطرق المعتمدة على أساليب الاتجاه الفراغي أو الأساليب اللغوية تحتاج إلى استفسار شامل أو كبير حتى تتمكن من إظهار قوتها الكاملة.

ومن الناحية العملية، تشير الإحصاءات إلى أن معظم الاستفسارات عادة ما تتكون من كلمة واحدة، ويكمن أحد أسباب قصر كثير من الاستفسارات في أن البحوث الأولى التي قام بها المستفيدون كانت تعتمد على أسلوب البحث البوليني الذي كان يثمر عن نتائج تضاهي الاستفسارات بالضبط، وبالتالي نادراً ما يتم الحصول على نتائج تضاهي الاستفسارات المطولة. أما الظاهرة الأخرى التي ارتبطت بنظم البحث المبكرة والتي شجعت على التعامل مع الاستفسارات القصيرة فهي أن أداء الحاسبات عادة ما كان يتدنى بصورة كبيرة عند التعامل مع الاستفسارات المطولة والمعقدة، وقد لوحظ أن كثيراً من المستفيدين كانوا حريصين على المحافظة على عاداتهم التي اكتسبوها عند تعاملهم مع تلك الحاسبات القديمة.

ومهما يكن من أمر، فقد ظل ميل المستفيدين إلى تقديم الاستفسارات القصيرة مسيطراً عليهم بدرجة كبيرة لا يمكن تبريرها بعوامل تاريخية، أو بالمساحات الضيقة لكتابة الاستفسارات التي تتوافر في تلك النظم القديمة، وهو نموذج كثيراً ما يتكرر في كل مكان تقريباً، بدرجة توحي بأن الاستفسارات المطولة تكاد تكون ممنوعة، وهناك ظاهرة أخرى غير صحية تتصل بالمستفيدين ـ لوحظ انتشارها على نطاق واسع، وهي أن القليل منهم من

يقرؤون أبسط إرشادات التعامل مع النظم. ومن الملاحظ الآن أن المكتبات الرقمية تفشل في تدريب روادها على تعليم إمكانات البحث بكفاءة، كما أن هؤلاء الرواد لا يقومون باستثمار كل المزايا المحتملة للنظم المتوافرة الآن.

التقويم :

ترتبط عملية تقويم الأداء ارتباطاً قوياً بعملية نظم استرجاع المعلومات. وهناك معياران مستخدمان في هذا الصدد منذ زمن طويل، أولهما: يعرف بالتحقيق precision أو الدقة (ويقصد به نسبة الإصابات الصالحة ـ أي إلى أي مدى تتفق المواد المسترجعة مع حاجة المستفيد كما عبر عنها في الاستفسار)، وثانيهما: يعرف بالاستدعاء recall (ويقصد به نسبة المواد الصالحة التي تم الوصول إليها، أي إلى أي مدى تم استرجاع جميع المواد الصالحة التي تتفق مع الاستفسار)، ومن الملاحظ أن كلا المعيارين يقيسان نتائج بحث واحد عن معلومات محددة. فلنفترض مثلاً أن هناك مجموعة مكونة من 10.000 وثيقة، خمسين وثيقة منها

تتناول موضوعاً محدداً، فمن المفترض أن البحث المثالي لما كتب عن هذا الموضوع يثمر عن استرجاع هذه الوثائق الخمسين دون غيرها من وثائق أخرى، لكن بحثاً حقيقياً أجري، وتم استرجاع خمس وعشرين وثيقة فقط من هذه الخمسين المرتبطة بالموضوع، و تبين أن هناك عشرين وثيقة من هذه الخمس والعشرين هي الصالحة أو المرتبطة بالموضوع، وأن خمس وثائق بعيدة عن الموضوع، فمعدل التحقيق في هذا البحث يحسب هكذا 20/25 = 0.8، كما يحسب معدل الاستدعاء هكذا 20/50 = 0.4.

ومما تجدر الإشارة إليه أن قياس معدل التحقيق أسهل بكثير من معدل الاستدعاء؛ حيث يتطلب حساب معدل التحقيق أن يقوم المستفيد بفحص الوثائق المسترجعة وتحديد عدد الوثائق الصالحة منها، ومن الواضح أن الأمر في المثال السابق لم يتطلب سوى فحص الوثائق الخمس والعشرين المسترجعة فقط وهو أمر سهل، أما بالنسبة لحساب معدل الاستدعاء، فما من سبيل للتعرف إلى جميع المواد الصالحة في مجموعة من الوثائق سوى فحص هذه المجموعة بأكملها للتثبت مما هو صالح أو غير صالح، وهي مهمة شاقة جداً.

ولقد اكتسب معيارا التحقيق والاستدعاء أهمية كبرى في مجال تطوير نظم استرجاع المعلومات، غير أنهما كانا

قد ظهرا عندما كانت الحاسبات بطيئة ومرتفعة الثمن بشكل أكبر مما هي عليه اليوم، بعد ذلك تكونت عملية الاسترجاع من بحث واحد لمجموعة كبيرة من البيانات، وكان النجاح أو الفشل حدثًا واحدًا. أما اليوم فقد أصبح البحث تفاعلياً، إذ يقوم المستفيد بصياغة الاستفسار، ثم يجري بحثاً أولياً، ويقوم بفحص نتائج هذا البحث، ثم يكرر العملية مع تعديل صياغة الاستفسار إذا لزم الأمر ذلك، وهو ما يعرف في اللغة الاصطلاحية لهذا المجال، بالبحث مع وجود الشخص في الحلقة searching with a human in the loop، وعليه فإن معياري التحقيق والاستدعاء في مثل هذه الجلسة أو العملية يجب أن يقاسا بنظرة إلى النتائج العامة، وليس بنتائج بحث واحد.

ومن الملاحظ أن معايير الأداء - كالتحقيق والاستدعاء - تتجه لقياس الخصائص الفنية لجميع النواحي المرتبطة بالنظم الآلية، ولا نقيس مطلقاً كيفية تفاعل المستفيد مع النظام، أو ما الذي يشكّل نتيجة دقيقة لبحث معين. وهناك في العديد من برامج البحث الحديثة جداً استراتيجية أو

طريقة لترتيب الإصابات المحتملة بشكل طبقي، وهذا ما يؤدي إلى إيجاد مستوى عالٍ من الاستدعاء على حساب العديد من الإصابات الخاطئة. وعليه فإن المعايير المطلوبة هي التي تقيس كفاءة الترتيب الطبقي في إعطاء المواد الأكثر صلاحية أو ارتباطاً بموضوع الرتب العليا.

ومما لا شك فيه أن المستفيدين يبحثون عن المعلومات لأسباب مختلفة، كما أنهم يلجؤون إلى أساليب مختلفة كذلك في بحثهم عن هذه المعلومات، فهم أحياناً يبحثون عن حقائق محددة، وأحياناً أخرى يبحثون بهدف استكشاف موضوع معين، ومن النادر ألا تواجههم في عملية البحث المشكلة العامة التي ترتبط بنظم استرجاع المعلومات، وهي الوصول إلى جميع المواد المناسبة تماماً لكل موضوع تم تحديده تحديداً دقيقاً، وبدون وجود مواد غير صالحة. ومع النظم الآلية التفاعلية، ولا يكتفي المستفيدون بإجراء بحث واحد، وإنما يقومون بتكرار سلسلة من الخطوات تتمثل في البحث المختلط والتصفح والتفسير وتنقية النتائج. وتعتمد كفاءة عملية استكشاف المعلومات على أهداف المستفيدين، وعلى مدى قدرة المكتبات الرقمية على تحقيق هذه الأهداف.

اللوحة رقم (10-7)

مشروع تبستر وتريك Tipster and TREC

تبستر مشروع طويل الأمد تبنته داربا، بهدف الارتقاء بجودة عمليات معالجة النصوص، وتركز اهتمامه على عدة قضايا، جميعها ذات أهمية كبرى بالنسبة للمكتبات الرقمية، وهي: التعرف إلى الوثائق أو اكتشافها document detection (الذي يجمع بين عملية استرجاع المعلومات من وثائق مخزنة وتحديد الوثائق المناسبة أو ذات العلاقة بموضوع البحث من بين كم كبير من النصوص الجديدة)، واستخراج المعلومات information extraction (أي تحديد مكان معلومات معينة داخل النص)، والتلخيص (أي تقليص أو ضغط حجم وثيقة معينة أو مجموعة وثائق).

وبمرور السنين تحول التركيز في مشروع تبستر من الاهتمام بعملية الاسترجاع القياسي للمعلومات إلى تطوير مكونات يمكن أن تؤدي مهام معينة، وتطوير بنية معمارية أو فنية تتقاسم فيها هذه المكونات مهامها، وكانت هذه المعمارية بمثابة جهد طموح سعى إلى الإفادة من مفاهيم البرمجة الموجهة للكائنات object-oriented programming في

تحديد فئات قياسية للمكونات الأساسية للمواد النصية، وعلى وجه التحديد: الوثائق والمجموعات والخصائص والشروح.

وفي أثناء المؤتمرات السنوية المهتمة بقضية استرجاع النصوص The annual Text Retrieval Conference المعروفة بتريك TREC، عرض باحثون أساليبهم الخاصة عن الأجسام القياسية للنصوص. وتعقد مؤتمرات تريك - وهي من الأمثلة على الأبحاث الكمية البارزة في مجال المكتبات الرقمية - تحت رعاية "المعهد الوطني للمعايير التقنية". the national institute of standard and technology وبمساعدة كثير من الهيئات الأخرى. فقد كوّن منظمو المؤتمرات مجموعة كبيرة من عدة ملايين من الوثائق النصية - تزيد على (5) جيجا من البيانات، ويقوم الباحثون بتقويم عملهم عن طريق محاولة تنفيذ مجموعة قياسية من المهام، وتتمثل إحدى هذه المهام في إجراء البحث في هذه المجموعة الكاملة عن موضوعات متنوعة قدمتها مجموعة مكونة من عشرين مستفيداً، وهناك اختبار آخر يجرى لتقييم النظم التي تقوم بالمضاهاة بين عدد كبير من الوثائق الواردة في مقابل مجموعة استفسارات قياسية، وكان من بين المشاركين في مؤتمرات تريك ممثلون من شركات تجارية كبرى، ومتعهدو نظم استرجاع معلومات صغار، ومجموعات بحثية من

الجامعات.

وتتيح مؤتمرات تريك فرصة للمقارنة بين مختلف الأساليب الفنية المهتمة بمثل تلك الطرق كالمكانز ووزن المصطلحات المعقد، وأساليب اللغة الطبيعية والتغذية المرتجعة لمدى الارتباط بالموضوع أو مدى الصلاحية، والتعليم المتقدم. وقد تم في السنوات الأخيرة تقويم مجالات أخرى في عملية استرجاع المعلومات، بما في ذلك الطرق التي سيتم استخدامها مع النصوص النصية، والوثائق الشفهية والاسترجاع متعدد اللغات، وهناك مسار آخر في المؤتمر تم تجربته على طرق تقويم الاسترجاع التفاعلي.

الفصل الحادي عشر
استكشاف المعلومات الموزعة أو اللامركزية

تعد عملية استكشاف معلومات مبثوثة وموزعة عبر العديد من النظم الآلية إحدى قوى التحدي التي تواجه عملية التشغيل المتداخل، فالمكتبات الرقمية على مستوى العالم تديرها جهات متعددة، وبأساليب إدارية متنوعة، وبتوجهات مختلفة، بناءً على طبيعة مجموعاتها، وعلى التقنيات التي يتم تطبيقها. كما أن قلة من هذه المكتبات هي التي تمتلك نسبة صغيرة من المواد التي يطلبها المستفيدون بكثرة؛ ومن ثم فإن المستفيدين يضطرون إلى طرق أبواب مختلف المجموعات والإفادة من الخدمات التي تقدمها كثير من المصادر المختلفة، وعليه كيف يستطيع المستفيد استكشاف المعلومات المبثوثة عبر هذه المصادر الكثيرة المحتملة، وكيف يمكنه الوصول إليها ؟

إن مصطلح التطبيقات الحاسوبية الموزعة أو اللامركزية Distributed Computing، مصطلح عام يعبر عن جميع الجوانب الفنية المرتبطة بعملية التنسيق بين الحاسبات الآلية التي تعمل مستقلة بعضها عن بعض حتى تتمكن من تقديم خدمة متناسقة. وتتطلب التطبيقات الحاسوبية الموزعة أن يكون هناك قدر مشترك من المعيارية الفنية أو التوحيد بين الحاسبات المختلفة، فقد يرغب المستفيد عند إجراء عملية البحث الموزعة - على سبيل المثال - أن يبحث في عدة مجموعات مستقلة بعضها عن بعض من خلال استفسار واحد، ثم يقوم بعد ذلك بمقارنة النتائج، ثم اختيار أكثر هذه النتائج قرباً من موضوع الاستفسار، ثم استرجاع المواد التي يقع عليها الاختيار من

بين هذه المجموعات. وعلاوة على ضرورة توافر صفة المعيارية الخاصة بعملية المشابكة، فإن هذه العملية - أي البحث الموزع - تتطلب توافر طريقة ما للتعرف إلى المقتنيات، ومعرفة ضوابط صياغة الاستفسارات وأساليب تقديمها، ومعرفة أساليب استرجاع النتائج، وطرق الحصول على المواد التي تم استكشافها، وقد تكون هذه المعايير رسمية معتمدة من قبل الجهات الرسمية المتخصصة في إصدار المواصفات القياسية، أو قد تكون معايير محلية طورتها مجموعة صغيرة من الجهات المتعاونة، أو مجرد اتفاقيات تنظم تداول منتجات تجارية معينة.

ولعل التوجه المناسب في هذا الصدد يتمثل في تطوير مجموعة شاملة من المعايير التي يمكن أن تتبناها جميع المكتبات الرقمية، ومع ذلك، فإن هذه الفكرة تخفق في معرفة تكاليف تبني هذه المعايير الشاملة والالتزام بها لاسيما في ظل هذه التغيرات المتسارعة التي نشهدها هذه الأيام؛ فالمكتبات الرقمية في تغير متواصل، وكل مكتبة تسعى لتطوير مجموعاتها وخدماتها ونظمها بحيث لا تتشابه مكتبتان في هذا التوجه، كما ينظر إلى أن عملية استبدال جزء من النظام أو تغييره لدعم معيار جديد هي عملية مضيعة للوقت، حيث يمكن أن تظهر نسخة جديدة من المعيار قبل أن تكتمل عملية الاستبدال هذه، أو ربما يكون الناس قد انصرفوا عن تطبيقه، واتجهوا نحو طريق آخر؛ فعملية التوحيد القياسي الكامل ما هي إلا سراب.

أما بالنسبة لعملية التشغيل المتداخل، فمن الواضح أن المكتبات الرقمية تواجه تحدياً كبيراً يتمثل في وجود نظم حاسوبية موزعة تعمل في سياق عالم تسوده حاسبات تعمل منفصلة بعضها عن بعض، وغير متوافقة من

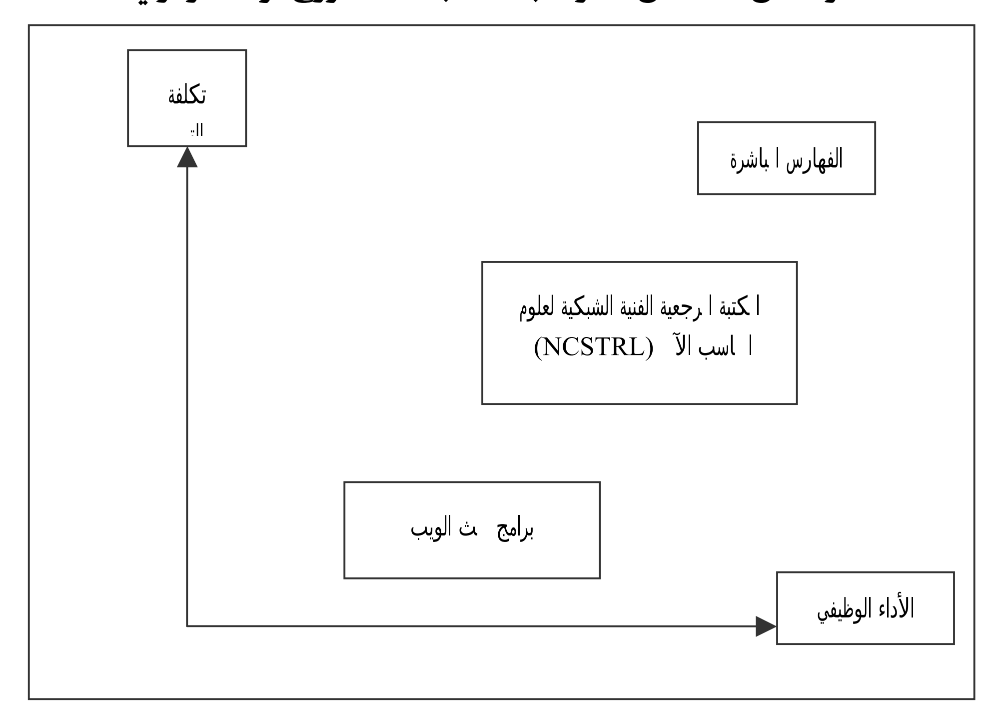
الناحية الفنية. ولكي يتسنى تبادل الرسائل فيما بين هذه الحاسبات، يتطلب الأمر وجود صيغ وبروتوكولات ونظم أمن تؤمن ذلك، كما أن الأمر يحتاج كذلك إلى وجود اتفاقيات تختص بدلالة الألفاظ للمساعدة في ترجمة تلك الرسائل وتفسيرها. ولكن يظل التحدي الأساسي متمثلاً في إيجاد طرق تحفز المكتبات الرقمية المنفصلة عن بعضها عن بعض على التعاون وعلى تضافر الجهود.

لا شك أن تبني طرق مشتركة يمنح المكتبات الرقمية مزيداً من القدرة على أداء وظائفها، غير أن لهذا النمط من التبني تكاليفه التي قد يكون بعضها مالياً كتكاليف شراء التجهيزات والبرمجيات، وتعيين الموظفين وتدريبهم؛ وقد تكون معظم هذه التكاليف في أحيان كثيرة مرتبطة بالنواحي التنظيمية، ونادراً ما يمكن تغيير أحد جوانب المكتبات الرقمية بمعزل عن الجوانب الأخرى. وإن طرح معيار جديد يستلزم إجراء تغييرات متداخلة في النظم الحالية، كما يتطلب كذلك تغييرات أخرى في سير العمل وتدفقه، وفي طبيعة العلاقات مع الموردين، وربما يتطلب تغييرات أخرى أكثر من ذلك.

ويصور الشكل رقم (١١-١) أحد النماذج التصورية التي يمكن الاستفادة منها عند التفكير في تدشين عملية التشغيل المتداخل، ويتم استخدام النموذج في هذه الحالة للمقارنة بين ثلاث طرق مختلفة من طرق البحث الموزع أو اللامركزي؛ حيث يشير المحور الأفقي في الشكل إلى الأداء الوظيفي الذي تقوم به الطرق المختلفة، في حين يشير المحور الرأسي إلى تكلفة تبني هذه الطرق. ويلاحظ أن الطريقة المثالية هي التي تتجسد في الأسفل يمين الشكل؛ حيث يتم تحقيق أعلى قدر من الأداء الوظيفي وبأقل تكلفة، فبرامج بحث الويب

تتمتع بقدرة متوسطة على أداء الوظائف، لكنها شائعة الاستخدام بسبب انخفاض تكاليف تبنيها، أما الفهارس المباشرة المعتمدة على معايير الفهرسة المقروءة آلياً المعروفة بمارك MARC، وعلى المعيار Z39.50، فإنها تتمتع بقدرة كبيرة على أداء الوظائف، لكنها محدودة الاستخدام بسبب معاييرها المعقدة، أما نظام المكتبة المرجعية الفنية الشبكية لعلوم الحاسب Networked Computer Science Technical Reference Library (NCSTRL) (المذكور في اللوحة ١١-٤) فيقع في مرحلة وسط بين هاتين الطريقتين.

الشكل رقم (١١-١)
مقارنة بين ثلاث من استراتيجيات البحث الموزع أو اللامركزي



وتقع الطرق التي يمكن استخدامها لأغراض عملية التشغيل المتداخل في

ثلاث فئات واسعة على النحو التالي:

- تتمتع معظم الطرق التي يشيع استخدامها على نطاق واسع لأغراض عملية التشغيل المتداخل اليوم بإمكانيات متوسطة على الأداء الوظيفي، وتكلفة تطبيق منخفضة. وتعد المعايير الأساسية المرتبطة بالويب المتمثلة في لغة ترميز النصوص الفائقة (HTML)، وبروتوكول نقل الملفات (HTP)، والمحدد الموحد لموقع المصدر (URL)، من النماذج التي تتمتع بهاتين الخاصتين. ومن الملاحظ أن البساطة التي اتسمت بها هذه المعايير قد أدت إلى تبنيها على نطاق واسع، غير أنها تحد من كفاءتها الوظيفية.

- توفر بعض الخدمات التي تستخدم أحدث التقنيات قدرة كبيرة على الأداء الوظيفي لكن تكلفة تبنيها عالية جداً، وخير مثال على ذلك معيار Z39.50، واللغة المعيارية الموحدة للترميز (SGML). وعادة ما يشيع استخدام هذه الأساليب في مجتمعات محددة، حيث تزداد أهمية الأداء الوظيفي، لكن تكلفة تبني هذه الطرق تعيق من تطبيقها في مجتمعات أكبر.

- تعد كثير من التطورات الحالية في مجال المكتبات الرقمية محاولات إيجاد أرضية مشتركة تتمثل في: زيادة القدرة على أداء الوظائف بتكلفة تطبيق متوسطة. وتشمل الأمثلة على ذلك كل من: معيار دبلن كور (Dublin Core)، ولغة الترميز الموسعة (XML)، ونظام اليونيكود (Unicode)؛ حيث حرص المصممون في جميع النماذج السابقة على توفير طرق معقولة التكلفة عند تبنيها؛ فدبلن كور مثلاً يسمح دائماً بأن تكون جميع الحقول اختيارية، كما أن نظام اليونيكود قدم صيغة يو تي إف-٨ التي تسمح بقبول البيانات الموجودة

في صيغة آسكي ASCII، كذلك تعد تكلفة تبني لغة الترميز الموسعة (XML) منخفضة نتيجة لارتباطها القوي بكل من لغة ترميز النصوص الفائقة (HTML)، واللغة المعيارية الموحدة الترميز (SGML).

وتجدر الإشارة إلى أن الشكل رقم (١١-١) ليس مقياساً محدداً، وما الأبعاد التي يتضمنها إلا مجرد أبعاد تصورية، تبين المبدأ الأساسي المتمثل في أن تكلفة تبني تقنية جديدة تعد أحد العوامل التي يجب أن تؤخذ بعين الاعتبار في جميع عمليات التشغيل المتداخل. كما أن العامل التقني يجب ألا ينظر إليه بمعزل عن العوامل الأخرى، وبدون وضع العامل التنظيمي في الاعتبار. فعندما يرغب منشئو إحدى المكتبات الرقمية في أن تعمل على نحو متبادل مع مكتبات أخرى، فإنهم غالباً يواجهون قضية المفاضلة بين البدائل المتاحة، ومن ثم اختيار أفضل الطرق المناسبة لمجتمعهم بشكل خاص، واختيار المعايير المقبولة بشكل جيد عموماً، والتي توفر لهم قدراً أقل من حيث الأداء الوظيفي. ولعل عملية اختيار الإصدارات الحديثة من البرمجيات تظهر مثل هذا التوتر، ففي الغالب توفر إحدى الإصدارات الجديدة للمكتبة الرقمية مزيداً من القدرة على أدائها الوظيفي، ولكن قلة من المستفيدين هم الذين يمكنهم الوصول إلى هذه المكتبات؛ حيث يستطيع منشئ أحد مواقع الويب - على سبيل المثال - أن يستخدم أكثر التيجان الأساسية للغة ترميز النصوص الفائقة، وأقل الصيغ الثابتة، والخدمات التي توفرها جميع إصدارات بروتوكول نقل الملفات الفائقة، في إنشاء هذا الموقع البسيط الذي يمكن لأي متصفح في العالم الوصول إليه ودعمه. وبدلاً من ذلك يمكن للمنشئ أن يختار أحدث نسخة من تقنيات الويب مع برامج تطبيقات جافا (جافا أبليت)، وصيغ

لغة ترميز النصوص الفائقة، وتجهيزات الأمن الداخلية، ونماذج عرض الصفحات (sheets)، وأدوات تحميل الصوت والصورة، وهي جميعها ستقدم أفضل خدمة فقط للمستفيدين الذين تتوافر لهم شبكات عالية السرعة مع أحدث المتصفحات، ومع ذلك فقد يرى آخرون أن جل هذه الخصائص غير مجدية بالنسبة لهم.

برامج بحث الويب :

تعد برامج الويب من أكثر النظم استخداماً لأغراض البحث الموزع أو اللامركزي، ومن أمثلة هذه البرامج إنفوسيك Infoseek، وليكوس Lycos، وألتافيستا AltaVista، وإكزايت Excite، وهي جمعيها نظم آلية تقوم بعملية تكشيف للمواد المتاحة على الإنترنت. وكما يظهر من الشكل رقم (١١-١) فإنها توفر مستوى متوسطًا من القدرة على الأداء الوظيفي، مع بعض الصعوبات القليلة في استخدامها والتي تتمثل في أمرين: أولهما أن مواقع الويب لا تتخذ أي إجراء خاص ليتم تكشيفها من قبل هذه البرامج، وثانيهما أنه ليس ثمة تكلفة يتحملها المستفيد سوى تكلفة الوقت الممل الذي يتحمله في مشاهدة الإعلانات التي تنتشر عبر صفحاتها، وليس ثمة من سبيل لجعل برامج بحث الويب شائعة الاستخدام إلى أكبر نطاق سوى الجمع بين رفع كفاءة الأداء الوظيفي، والتحرر من صعوبات الاستخدام.

ومن الملاحظ أن الغالبية العظمى من هذه البرامج تشترك في بنيتها الفنية الأساسية، وإن كانت ثمة اختلافات كثيرة بينها في بعض التفاصيل، ولا يستثنى من ذلك سوى برنامج "ياهو"، الذي يعتمد في تنظيمه إلى عملية التصنيف الموضوعي. أما النظم الأخرى فتتكون من جزأين رئيسين: يعرف

أولهما بزاحف الويب أو الويب كرولر Web Crawler، (وهو الذي يقوم ببناء كشاف للمواد المتاحة على الإنترنت)، ويعرف ثانيهما بمحرك الاسترجاع research engine، (وهو الذي يسمح للمستفيدين بالبحث في هذا الكشاف).

زواحف الويب (ويب كرولرز) :

زاحف الويب هو برنامج تكشيف يتتبع الروابط الفائقة باستمرار، ويقوم بتجميع قائمة بالصفحات التي يعثر عليها، ويقوم كذلك ببناء كشاف مستمر النمو من صفحات الويب، وذلك من خلال تكرار الخطوات الأساسية القليلة التي يقوم بها. ويحتفظ الزاحف كذلك داخله بقائمة بالمحددات الموحدة لمواقع المصادر URLs المعروفة للنظام، سواء أكان قد تم تكشيف الصفحات المقابلة لهذه المحددات أم لم يتم تكشيفها بعد، ثم يقوم بعد ذلك باختيار المحدد الموحد الخاص بأي صفحة من صفحات الويب المرمزة عن طريق لغة ترميز النصوص الفائقة والتي لم يقم بتكشيفها من قبل، على أن يقوم باسترجاع هذه الصفحة ويعيدها مرة ثانية إلى النظام الآلي المركزي من أجل تحليلها، حيث يقوم هناك أحد برامج التكشيف الآلي بفحص الصفحة وإنشاء تسجيلة تكشيف لها، تضاف بعد ذلك إلى الكشاف الشامل، كما يقوم بالتقاط الروابط الفائقة الموجودة في هذه الصفحة والتي تشير إلى صفحات أخرى، بحيث يتم إضافة الروابط الفائقة الجديدة إلى قائمة المحددات الموحدة الموجودة لديه، وذلك من أجل استكشافها مستقبلاً أو في مرات لاحقة.

وتجدر الإشارة إلى أن هناك كثيراً من الاختلافات والمشكلات الفنية الدقيقة التي تكمن وراء هذا الإطار المبسط الذي تم تصويره، وتتمثل إحدى

التساؤلات عن المحدد الموحد الذي يجب على الزاحف أن يزوره بعد ذلك، ففي أي لحظة قد يكون هناك لدى زاحف الويب ملايين المحددات الموحدة التي يتم استكشافها وغير المتوقعة، لكن لا تتوافر له سوى معلومات قليلة يمكن الاسترشاد بها في اختيار المحدد التالي. وتشتمل معايير الاختيار على عدة تساؤلات أو عوامل منها: ما مدى حداثة هذا المحدد ؟ وما عدد المحددات الأخرى المرتبطة به ؟ وهل هذا المحدد يمثل إحدى الصفحات الرئيسية أم إحدى الصفحات الفرعية داخل مجموعة هرمية من الصفحات ؟

وبالرغم من كل ذلك، فلا تزال المشكلة الكبرى هي عملية التكشيف، فالزواحف تعتمد أساساً على عملية التكشيف الآلي - التي نوقشت في الفصل العاشر - في إنشائها للتسجيلات التي تقدم للمستفيدين، وبذلك فإن هذه الزواحف تواجه قضايا التكشيف الآلي الأساسية، فهناك ملايين من الصفحات التي يتم إنشاؤها من قبل أشخاص ذوي تصورات متباينة عن كيفية بناء المعلومات. كما أن الصفحات التي تعد مثالية لا تقدم إلا القليل من المؤشرات اللازمة لأغراض التكشيف الآلي، بل إن بعض منشئي الصفحات وبعض الناشرين يتعمدون التضليل؛ فيملؤون صفحاتهم بالمصطلحات التي يرجحون أن يستخدمها المستفيدون أملاً في أن تحظى صفحاتهم بالرتب الأعلى في سلم ترتيب قوائم الصفحات، والذي عادة ما يبنى على استفسارات البحث الشائعة. وأخيراً يمكن القول إنه بدون صفحات مبنية بناءً محكماً، أو ما وراء بيانات منتظمة، لا يمكن للتسجيلات الكشفية أن تحقق أعلى درجات الدقة عند عمليات البحث، بيد أن هذه التسجيلات يمكن أن تكون مناسبة لعمليات الاسترجاع المبسطة.

البحث في أحد كشافات صفحات الويب :

تسمح برامج بحث الويب للمستفيدين بالبحث في كشافاتها من خلال استخدام طرق استرجاع المعلومات التي أوضحناها في الفصل العاشر. وعادة ما تكون هذه الكشافات المجهزة بأسلوب يكفل لأعداد كبيرة من المستفيدين إمكانية البحث فيها بكفاءة وفي آن واحد. ونظراً لأن التسجيلات التي تتضمنها الكشافات قد لا تتمتع بالقدر الكافي من الدقة، بالإضافة إلى أن المستفيدين قد يفتقرون إلى التدريب الكافي للتعامل مع هذه السجلات، أو مع أساليب البحث، فإن برامج البحث تقوم باتباع استراتيجية تحديد جميع التسجيلات التي تضاهي الاستفسار، حتى وإن كانت هذه المضاهاة غير دقيقة، ثم تقوم بعرضها للمستفيد مرتبة طبقياً إلى حد ما.

ولعل معظم مستخدمي برامج الويب يتفقون على أنها برامج جيدة إلا أن ثمة كثير من الصعوبات الواضحة التي تكتنف استخدامها، فخوارزميات الترتيب الطبقي تفتقر إلى المعلومات الكافية التي يجب أن تبنى عليها قرارات الترتيب. ونتيجة لذلك، فإن هذه البرامج قد تعطي بعض الصفحات الهامشية رتباً عالية لا تستحقها، ومن ثم ترد المواد الهامة في ذيل القائمة، وترد المواد الأقل أهمية أو التافهة على رأس القائمة، يضاف إلى ذلك أن برامج التكشيف يصعب عليها التعرف إلى المواد المكررة، مع أنها تحاول القيام بتجميع المواد المتشابهة، وبما أن هذه المواد المتشابهة يتوقع أن ترد جميعها في رتبة واحدة. فإن برامج التكشيف غالباً ما تقوم بإعداد قوائم مطولة من المواد المتماثلة تقريباً. ومن إحدى الطرق الجيدة للترتيب الطبقي

تلك الطريقة التي يتبعها محرك البحث جوجل Google، والتي تتمثل في قيامه بعمليات إحصائية للروابط. ويتميز جوجل بفاعلية جيدة في إيجاد المواد التمهيدية أو العامة حول الموضوع الذي يبحث عنه.

وعلى الرغم من أن زواحف الويب تقوم باستكشاف الويب بشكل مستمر بحيث يتم لها في النهاية العثور على كل شيء تقريباً، يلاحظ أن المواد الهامة قد لا يتم تكشيفها إلا بعد مضي عدة شهور من تحميلها على الويب، وعلى العكس من ذلك، يلاحظ أن البرامج لا تقوم بأداء وظيفتها على نحو جيد، والتي تتمثل في العودة مرة أخرى للتثبت من مدى سحب المواد أو استبعادها من على الويب، ولذلك نجد كثيراً من مداخل الكشافات تشير إلى مواد لم يعد لها وجود على الشبكة، أو أنها انتقلت من مواقعها إلى مواقع أخرى.

وهناك تهديد آخر يحيط بمدى كفاءة عملية تكشيف الويب يتمثل في أن زاحف الويب لا تستطيع تكشيف المواد التي لا يمكن الوصول إليها بشكل مباشر، فإذا كانت إحدى صفحات الويب خاضعة للحماية عن طريق إحدى صور المصادقة أو ضمانات الثقة authentication، أو إذا كانت إحدى صفحات الويب عبارة عن واجهة تفاعل لقاعدة بيانات، أو لمجموعات إحدى المكتبات الرقمية، فإن برامج التكشيف لا تتمكن من معرفة أي شيء عن المصادر الموجودة وراء الواجهة. وبما أن هناك الكثير والكثير من صفحات الويب عبارة عن واجهات تفاعل تتحكم فيها برامج جافا أو غيرها من برامج النصوص الأخرى، فإن برامج التكشيف تفقد كثيراً من المعلومات الهامة.

ومع أن نقاط الضعف هذه تبدو جوهرية، فإنه لا ينبغي المبالغة في ذلك، والدليل على ذلك يظهر من خلال الممارسة العملية؛ حيث يمكن لمستخدمي الويب الذين يتمتعون بقدر قليل من الخبرة، والمعتمدين في استخدامهم للويب على مجموعة من الأدوات المتنوعة، والتي تكون في الغالب خدمات بحث الويب -أن يصلوا للمعلومات التي يبحثون عنها. ومع أن هذه البرامج ليست مبرأة من كل نقص، فإنها تعد جيدة إلى حد بعيد، والأهم من ذلك أن استخدامها متاح مجاناً.

وتجدر الإشارة إلى أن معظم خدمات البحث كان لها جذورها في ما يعرف بمجموعات البحث research groups، إلا أنها ما فتأت أن أصبحت -بسرعة -شركات تجارية، ولقد كان لعدم فرض رسوم مالية على الخدمات الأساسية أثر كبير على كل من الإنترنت من ناحية، وعلى الشركات نفسها من ناحية أخرى؛ فقد أدى سعي هذه الشركات إلى تحصيل إيرادات مالية إلى التنافس المحموم في جلب الإعلانات، كما أن هذه الشركات اتجهت بسرعة إلى دخول الأسواق عبر وسائل معينة مثل منح تراخيص برمجياتها لشركات أخرى ترغب في بناء كشافات لمواقعها على الويب.

أما الجانب غير المرغوب في هذا النموذج التجاري فيتمثل في أنه يحد من الدافع إلى إنشاء كشاف شامل، فقد كانت برامج التكشيف تهدف في البداية إلى تكشيف الويب تكشيفاً شاملاً، ولكن في ظل النمو الذي تشهده الويب، وبحكم أن إدارة برامج البحث تحولت إلى مشروعات تجارية، فقد أصبحت الشمولية المنشودة في مرتبة ثانوية تالية للتحسينات في واجهات التفاعل وفي الخدمات المساعدة. ولا شك أن إنشاء كشاف جيد أو دقيق

ومحدث بمعنى الكلمة، يتطلب استثمارات ضخمة، وتقتنع معظم الشركات بأداء وظيفتها على نحو معقول، ولكن إذا زادت دوافعها، فسوف تكون كشافاتها أفضل مما هي عليه الآن.

اللوحة رقم (١١-١)

رُتب الصفحات وجوجل **Page Ranks and Google**

يعد تحليل الاستشهادات المرجعية من الأدوات الشائعة الاستخدام في العلوم، فالمقالات التي تستشهد بمقالات أخرى عادة ما تكون جميعها مرتبطة بعضها ببعض ارتباطاً موضوعياً. كما أن المقالات التي يكثر الاستشهاد بها عادة ما تكون أكثر أهمية من تلك المقالات التي لا يستشهد بها على الإطلاق. وقد قام كل من لورانس بيدج Lawrence Page، وسيرجي براين Sergy Brain وزملاؤهما في جامعة ستانفورد، بتطبيق فكرة هذا الأسلوب على الويب، مستخدمين الأنماط التي تتخذها الروابط الفائقة فيما بين الصفحات كأساس للترتيب الطبقي للصفحات، وطوروا برنامجاً تجريبياً لبحث الويب، يعرف باسم "جوجل".

فتخيل - على سبيل المثال - أنه باستخدام برامج بحث مختلفة للبحث عن كل ما يتصل "بجامعة ستانفورد"، وتبين أن ثمة أكثر من ٢٠٠,٠٠٠ صفحة تتصل بالتساؤل عن كلمة "ستانفورد Stanford"، ويلاحظ أن معظم برامج البحث يصعب عليها التمييز بين الصفحات ذات الطبيعة المحلية أو ذات القيمة الهامشية، وتلك الصفحات ذات الأهمية الكبرى فجميع برامج بحث الويب تُوجد أعدادًا هائلة من الصفحات التي تضاهي ذلك الاستفسار؛ أي ما يتصل بجامعة ستانفورد. ولكنها في معظم الحالات تخفق في ترتيب هذه الصفحات

ترتيباً طبقياً يعتبره المستفيدون دقيقاً بالشكل الذي يكفل للصفحات المهمة أن ترد في رأس القائمة.

وعندما قدم التساؤل عن جامعة ستانفورد لنظام جوجل لبحث الويب جاءت النتائج العشرة الأولى على النحو التالي:

- Stanford University Home page (www.stanford.edu)
- Stanford University Medical center (www.med.stanford.edu)
- Stanford University Library and information Resources (www.sul.stanford.edu/)
- Stanford Law school (www.leland.stanford.edu/group/law/)
- Stanford Graduate school of Business (www.gsb.stanford.edu/)
- Stanford University school of earth science (pangea.stanford.edu/)
- SUL: copyright & fair use (fairuse.stanford.edu/)
- Computer Graphic at Stanford University (www.graphic.stanford.edu)
- SUMMIT (Stanford University) Home page (summit. stanford.edu).
- Stanford medical Informatics (camis.standord.edu/).

وقد اتفقت الغالبية العظمى من المستفيدين على أن هذه القائمة جيدة.

والطريقة الأساسية التي استخدمها نظام "جوجل " لترتيب هذه الصفحات طريقة مبسطة، تتمثل في اعتمادها على قاعدة مفادها "أن الصفحات التي يتواتر الإشارة إليها من قبل الصفحات في شكل روابط فائقة تشير إلى صفحات أخرى، ترد في رتبة أعلى من الصفحات التي تقل

الإشارة إليها، كما أن الروابط التي وجدت في الصفحات التي احتلت رتباً أعلى تأخذ وزناً أكبر من الروابط التي وجدت في الصفحات التي احتلت رتباً متدنية. وبما أن صفحات الويب حول العالم تقوم بعمل روابط إلى الصفحة الرئيسية لكلية الحقوق بجامعة ستانفورد، فإن هذه الصفحة تقوم بدورها بعمل روابط إلى العديد من الصفحات الأخرى، مثل الصفحة الرئيسية للجامعة، ومن ثم فإن هذه الأخيرة جاءت في رتبة متقدمة بسبب الإشارة إليها من قبل صفحة احتلت رتبة متقدمة كذلك.

ويتطلب حساب رتب الصفحات عملية حسابية ذكية، ولكي تفهم الفكرة الأساسية لهذه العملية، تخيل أن أمامك مصفوفة كبيرة تضم كل صفحة موجودة على الإنترنت، وأمام كل صفحة قائمة بالروابط التي تشير إليها، ففي بداية الأمر، يتم ترتيب الصفحات على نحو متساوٍ، ثم يتم بعد ذلك حساب عملية الترتيب الجديدة لهذه الصفحات على أساس عدد الروابط الموجودة أمام كل صفحة، على أن يتم عمل وزن لكل صفحة بناء على رتبة الصفحات التي وردت بها الروابط لتلك الصفحة linking pages، منسوبة إلى عدد الروابط المأخوذة من كل صفحة. ثم يتم بعد ذلك استخدام هذه الرتب لتكرار آخر، وتستمر العملية على هذا النحو حتى تنتهي العملية الحسابية.

وتعد العملية الحسابية الآلية computation الحقيقية تحسينًا لهذه الطريقة، ففي عام ١٩٨٨م كان جوجل يضم ما يقرب من ٢٥ مليون صفحة، تم اختيارها عن طريق عملية تم أخذها من رتب الصفحات التي أشارت أو قدمت ربطاً لهذه الصفحات. وتتوافر للبرنامج عوامل وزن لحساب الصفحات التي

ليس لها روابط، ولحساب مجموعات الصفحات التي بها روابط لصفحة واحدة أخرى فقط. ويرفض البرنامج الصفحات التي يتم إنشاؤها بصورة آلية أو ديناميكية من قبل نصوص واجهات بوابات العبور الموحدة Common Gateway Interface (CGI). ذلك لأن النظام كان قادراً على جمع هذه الصفحات وتكشيفها وترتيبها في خمسة أيام باستخدام حاسبات محطات العمل القياسية فقط، ولكن ليبرهن على قوة حاسبات اليوم.

إن استخدام الروابط لإنشاء رتب الصفحات يساعد على حل مشكلتين من المشكلات التي تضعف عمل برامج بحث الويب، وهاتان المشكلتان تتمثلان في التساؤلين التاليين: بما أن البرامج لا تستطيع تكشيف كل صفحة متاحة على الويب في الوقت نفسه الذي تظهر فيه، فما الصفحات التي يجب أن تكشف قبل غيرها؟ وكيف تتمكن البرامج من ترتيب الصفحات التي تسترجع عند طرح التساؤلات البسيطة حتى يتم إعطاء الأولوية للصفحات الأكثر أهمية ؟

المكتبات الرقمية الاتحادية :

إن تخفيف حدة الجدل القائم حول قضيتي الأداء الوظيفي وتكلفة تبني تقنية معينة يعتمد على الظروف التي يتم في سياقها هذا الجدل، فقد يكون من المناسب في بعض الأحيان اختيار تقنية بسيطة، والسعي الحثيث لتحقيق معدلات تغطية واسعة غير متعمقة من التشغيل المتداخل، وقد يكون من الحكمة في أحيان أخرى أن يتم اختيار تقنية تتمتع بقدر عال من الكفاءة في الأداء الوظيفي، ولكن ذلك يتطلب تحمل تكاليف عالية، ولا يقدر على تبني

هذه الأساليب المكلفة سوى المكتبات الرقمية ذات الطموحات الكبيرة، وإن كانت سوف تحقق مستوى عاليًا في أدائها الوظيفي.

ومصطلح المكتبة الرقمية الاتحادية يستخدم ليشير إلى مجموعة من المنظمات التي تعمل بعضها مع بعض بشكل رسمي أو غير رسمي، متفقة فيما بينها على دعم مجموعة من الخدمات والمعايير المشتركة، ومن ثم توفير إمكانية التشغيل المتداخل بين أعضائها. وتجدر الإشارة إلى أن أعضاء هذا الاتحاد يمكن أن يكون لكل منهم نظام يختلف عن النظم الأخرى اختلافاً تاماً، إلا أن ذلك لا يجسد مشكلة ما دامت جميعها تدعم مجموعة من الخدمات المتفق عليها. ويتطلب الأمر اتفاقهم على كل من المعايير الفنية والسياسات، بما في ذلك الاتفاقيات المالية وتشريعات الملكية الفكرية، وسياسات الأمن والخصوصية.

ويقدم بحث أجري في جامعة إلينويز في أربانا شامبين University of Illinois at Urbana Champaign مثالاً للصعوبات التي تكتنف عملية التشغيل المتداخل، ففي الفترة من ١٩٩٤- ١٩٩٨م وكجزء من مبادرة المكتبات الرقمية، بدأ فريق من العاملين في مكتبة جرينجر الهندسية Grainger Engineering Library بإنشاء مكتبة اتحادية لمقالات الدوريات المنشورة من قبل العديد من الناشرين العلميين البارزين. وبما أن كل ناشر من هؤلاء كان قد خطط لإتاحة دورياته بصيغة اللغة المعيارية الموحدة للترميز (SGML)، فقد بدا ذلك وكأنه فرصة لإنشاء ذلكم الاتحاد. ولأنه كان من المفترض على الجامعة أن تقدم خدمات مركزية، كخدمات البحث، فقد كان على الناشرين تأمين تلك الدوريات. لكن عدم تناغم الأساليب التي استخدم بها الناشرون

اللغة المعيارية الموحدة للترميز؛ قد سبب مشكلات في هذا الصدد، حيث كان لكل ناشر من هؤلاء الناشرين تعريفه الخاص بشكل الوثيقة. وقد اضطرت الجامعة إلى القيام بمهام مطولة للتوفيق بين الدلالات اللفظية لمعرفات أنواع الوثائق the semantics DTDs، من أجل إتاحة إمكانية اشتقاق المعلومات الكشفية، ومن أجل تطوير واجهة متماسكة للمستخدمين. وقد ثبت أن ذلك كان عملاً مضنياً مما حدا بالجامعة أن تلجأ إلى استنساخ المعلومات من حاسبات الناشرين وتحميلها على نظام واحد، ثم تحويلها إلى معرف واحد من معرفات أنواع الوثائق. وإذا كان قد قدر لمجموعة بحث في جامعة مرموقة أن تواجه هذه الصعوبات مع جزء متناغم نسبياً من المعلومات، فليس من المدهش أن يواجه الآخرون مثل تلك الصعوبات.

اللوحة رقم (١١-٢)

مكتبة جامعة إلينويز الاتحادية للإنتاج الفكري في العلوم

The University of Illinois Federated Library of Scientific Literature

تعد مكتبة جرنجر الهندسية (The Grainger Engineering Library) في جامعة إلينويز مركزاً لأحد النماذج الأولية prototype لمكتبة اتحادية لمقالات الدوريات العلمية. وقد بدأ العمل فيها كجزء من مبادرة المكتبات الرقمية بقيادة كل من بروس تشاتز Brauce Schatz ووليم ميسكو William Mischo. وبحلول عام ١٩٩٨م وصلت مجموعة المقتنيات التي خضعت للاختيار إلى ٥٠,٠٠٠ مقالة من المقالات المنشورة في دوريات كل من معهد المهندسين الكهربائيين والإلكترونيين (IEEE)، وجمعية الحاسب في هذا المعهد، والجمعية الفيزيائية الأمريكية، وجمعية المهندسين المدنيين

الأمريكيين، والمعهد الأمريكي لعلوم الطيران والفضاء. وقد اتفق على أن تقوم كل جهة من هذه الجهات بتقديم مقالات من دورياتها العلمية مهيأة باللغة المعيارية الموحدة للترميز في الوقت نفسه الذي تنشر فيه تلك الدوريات مطبوعة.

وقد طبق هذا النموذج الأولي مفاهيم استرجاع المعلومات من النص المرمز marked-up text والتي نوقشت كثيراً ولم تطبق إلا قليلاً. وقد بدأت المرحلة الأولى (التي اتسمت بالصعوبة والتعقيد) بعملية التوفيق بين معرفات أنواع الوثائق التي استخدمها ناشرو الدوريات، حيث كان كل منهم يستخدم معرفاً خاصاً به لعرض العناصر البنائية لوثائقه. وقد تجسدت بعض أوجه الخلاف في الجوانب الدلالات اللفظية، حيث كان يعبر عن عنصر المؤلف في بعض الأحيان - على سبيل المثال - بالتاج <author>، في حين يعبر عنه في أحياناً أخرى بالتاج <aut>، وفي أحيان ثالثة بالتاج <au>. وكانت الخلافات الأخرى تعكس أوجه اختلاف دلالية جوهرية. ومن أجل تنفيذ عمليتي التكشيف والاسترجاع، لجأ القائمون على المشروع إلى وضع برمجية تقوم بعمل تخطيط لتيجان كل معرف من معرفات نوع الوثائق على هيئة صيغة مقبولة من الجميع (canonical set) على أن تقوم واجهات التفاعل مع المكتبة الرقمية باستخدام هذه التيجان، بحيث يستطيع المستفيد أن يبحث عن نص في سياق معين، كما هو الحال مع دليل الصور أو الأشكال figure legend. وفي مخطط الشكل رقم (١١-١) يقع استخدام هذه المجموعة من التيجان في أعلى اليسار. ومع أن عملية تحويل الترميز التي أعدت لجميع مفردات هذه

المقتنيات إلى هذه الصيغة تتطلب تكلفة عالية كلما أضيفت مجموعة مقتنيات جديدة إلى الاتحاد، فإنه يهدف إلى توفير مستوى عال من الأداء الوظيفي.

وبسبب الصعوبات الفنية، تمثلت الخطوة الأولى من التنفيذ في تحميل جميع الوثائق في مستودع واحد في جامعة إلينويز، على أن تدعو الخطط المستقبلية إلى استخدام المستودعات التي تم إعدادها من قبل الناشرين. وَثَمَّ اهتمام آخر يقضي بتوسعة مجال المقتنيات لتشمل قواعد البيانات الببليوجرافية، وفهارس، وكشافات أخرى.

وبالرغم من أنه قد ثبت أن عملية التنفيذ الأولى لم تكن أرضاً قوية لدراسة المستفيدين ورغباتهم، فإن توفيرها لأساليب إضافية قوية لإجراء عمليات من جانب المستفيدين كان أمراً محموداً، ولكنه شجع على تلقي الاستفسارات. وقد أشار المستفيدون إلى أن الأشكال البيانية والمعادلات الرياضية غالباً ما تكون موضحة بشكل كبير في محتوى المقالات أكثر من مستخلصات تلك المقالات والفقرات الختامية الخاصة بالنتائج التي تتضمنها تلك المقالات. كما أظهرت هذه التجارب - مرة أخرى - أن المستفيدين يواجهون صعوبات كبيرة في العثور على الكلمات الصحيحة التي يجب استعمالها في استراتيجيات البحث عندما لا يكون هناك ضبط للمصطلحات المستخدمة في البحوث وفي مستخلصاتها، وفي نظام البحث.

الفهارس المباشرة ومعيار Z39.50 :

لكثير من المكتبات فهارس مباشرة لمقتنياتها التي يتاح الوصول إليها مجاناً من خلال الإنترنت، ويمكن اعتبار هذه الفهارس شكلاً من أشكال

الاتحادات. وعادة ما تتبع قواعد فهرسة الأنجلو الأمريكية Anglo-American Cataloging Rules في إعداد تسجيلات هذه الفهارس من خلال استخدام صيغة الفهرسة المقروءة آلياً المعروفة بمارك، وتقوم المكتبات بتقاسم عملية إعداد هذه التسجيلات فيما بينها بهدف تخفيض التكلفة. وقد قام المكتبيون بتطوير معيار Z39.50 (الذي سيرد وصفه في اللوحة رقم ١١-٣)، من أجل تلبية احتياجات ما يتصل بكل من تقاسم إعداد التسجيلات، وعمليات البحث اللامركزي. وفي الولايات المتحدة نشطت كل من مكتبة الكونجرس ومركز الحاسب الآلي للمكتبات على الخط المباشر OCLC، ومجموعة مكتبات البحث في تطوير هذا المعيار ونشره، وقد كان هناك العديد من التطبيقات المستقلة لمعيار Z39.50 في المواقع الأكاديمية، والموردين التجاريين. كما تم استيعاب التكلفة العالية للانضمام لمثل هذا الاتحاد عبر العقود الماضية، وتم تعويض هذه التكاليف من خلال الوفورات المالية التي تحققت من جراء عمليات الفهرسة المشتركة.

وتتمثل إحدى جوانب التطبيقات الأساسية لمعيار Z39.50 في مجال الاتصالات بين الحاسبات الخادمة. ويمكن لنظام الفهرسة في إحدى المكتبات الكبرى أن يستخدم هذا المعيار لبحث مجموعة من الفهارس المتشابهة، ليرى ما إذا كان أي منها يوجد لديه نسخة من عمل ما، أو قامت بإعداد تسجيلة له. ويمكن للمستفيدين النهائيين أن يستخدموا حاسباً عميلاً واحداً يدعم معيار Z39.50 لبحث العديد من الفهارس، سواء تم ذلك على التوالي أو على التوازي. وقد جنت المكتبات وروادها فوائد جمة من خلال عمليات تقاسم الفهارس بهذه الطرق، وبالرغم من ذلك، لا تزال

عملية التشغيل المتداخل بين الفهارس المباشرة public access غير مكتملة الأركان. ومع أن بعض تطبيقات معيار Z39.50 تمتع ببعض المزايا التي لا تتمتع بها غيرها، فإن السبب الأساسي لعدم الاكتمال يتمثل في أن الفهارس الفردية يتم إعدادها من قبل أشخاص يعملون وفق احتياجات مجتمعاتهم المحلية، ولم يكن دعم المؤسسات الأخرى يحظى بالأولوية على الإطلاق. وحتى عندما تتقاسم المؤسسات نسخاً متوافقة من البروتوكول Z39.50، فإن الاختلافات في كيفية تنظيم الفهارس وإتاحتها للعالم الخارجي ستظل باقية.

اللوحة رقم (١١-٣)

معيار **Z39.50**

في رحاب مجتمع المكتبات تم تطوير المعيار المعروف بـ Z39.50، والذي يسمح لأحد الحاسبات (العميل) بالقيام بعمليات البحث والاسترجاع للمعلومات التي توجد لدى حاسب آخر (خادم قاعدة البيانات)، ويعد هذا المعيار هامًّا بحكم بنيته الفنية، وبسبب شيوع استخدامه في نظم المكتبات. ومع أن هذا البروتوكول لا يعد- من الناحية النظرية - مخصصاً للتعامل مع أشكال محددة من المعلومات، أو نوعية معينة من قواعد البيانات، فإن معظم توجهات تطويره تركزت على البيانات الببليوجرافية، كما تركزت معظم تطبيقاته على البحوث التي تستخدم السمات الببليوجرافية في بحث قواعد بيانات تسجيلات مارك، واسترجاع هذه التسجيلات إلى الحاسب العميل.

ويبنى المعيار Z39.50 على وجهة نظر مجردة للبحث في قواعد البيانات، حيث يفترض أن الحاسب الخادم يقوم باختزان مجموعة من قواعد البيانات مع كشافات قابلة للبحث. وتبنى التفاعلات على فكرة أو مفهوم

"الجلسة session"، إذ يقوم الحاسب العميل بالاتصال بالحاسب الخادم، ثم يقوم بإجراء سلسلة من التفاعلات، ثم يقوم بإنهاء الاتصال بعد ذلك. وأثناء وقت الجلسة يقوم كل من الخادم والعميل بتذكر حالة تفاعلاتهما، ومن المهم أن نعي أن العميل هنا هو جهاز حاسب آلي، وأن تطبيقات المستخدم النهائي للبروتوكول Z39.50 تستلزم واجهة للمستخدمين تستخدم لأغراض الاتصال بالمستفيد، ولا يقوم البروتوكول بطلب بيانات عن شكل واجهة المستفيد، ولا عن كيفية اتصالها بالحاسب العميل الذي يدعم بروتوكول Z39.50.

وتبدأ الجلسة الفعلية بقيام العميل بإجراء الاتصال بالخادم، ومن ثم تبادل المعلومات الأولية بينهما، مستخدماً في ذلك وسيلة "خاصية init"، ويقوم هذا التبادل الأولي بعمل اتفاقية على الأمور الأساسية، مثل الحجم المفضل للرسالة، كما يمكن أن تحتوي كذلك على إجراءات "المصادقة authentication " وإن كانت عمليات المصادقة الحقيقية تتم خارج نطاق المعيار. ويمكن للعميل بعد ذلك استخدام خدمة "التفسير explain" ليستفسر من الخادم عن كل من قواعد البيانات المتاحة للبحث، وعن الحقول القابلة للبحث فيها، وعن تركيب كلمات البحث وجمله syntax، وعن أشكال النص التي يدعمها، إلى غير ذلك من خيارات.

وتسمح "خدمة البحث search service" للعميل بأن يقدم استفسارًا إلى قاعدة البيانات كما في المثال التالي :

"في قاعدة البيانات المسماة "Books" ابحث عن جميع التسجيلات التي تشتمل حقول العنوان فيها على القيمة "Evangeline"، والتي تشتمل حقول بيانات التأليف فيها على القيمة "Longfellow".

ومع أن المعيار يقدم عدة بدائل لتركيب الكلمات وجمل البحث[١] التي تستخدم لتحديد مجال البحوث، فإن الاستفسارات البولينية[٢] هي التي يتم تنفيذها بشكل شائع، حيث يقوم بتنفيذ البحث، واسترجاع النتائج. ومن الإمكانيات المتميزة للبروتوكول Z39.50 تمكن الخادم من حفظ نتائج البحث، بحيث يمكن إحالة العميل في استفسار لاحق إلى هذه النتائج. وهكذا يستطيع العميل تكوين مجموعة ضخمة من النتائج الناجمة من خلال الاستفسارات الدقيقة، كما يمكنه طلب عرض أي تسجيلة من هذه المجموعة، بدون القيام بالبحث الكامل لقاعدة البيانات مرة أخرى.

وبناء على خصائص استفسارات البحث، يمكن أن تسترجع تسجيلة واحدة أو أكثر للعميل، كما يتيح المعيار عدة أساليب متنوعة للعملاء تمكنهم من معالجة نتائج البحث، مثل إمكانات الفرز والحذف، وعندما تنتهي عملية البحث، يحتمل أن يبدأ العميل بإرسال استفسار جديد. وهذه التساؤلات تجعل الخادم يقوم بإرسال تسجيلات محددة من مجموعة النتائج إلى العميل في شكل محدد. وتحتوي الخدمة الحالية على عدة خيارات تتصل بعملية التحكم

(١) تعرف باستراتيجيات البحث (المترجمان).

(٢) تعرف بالروابط البولينية أو المنطقية (المترجمان).

في المحتوى وفي الصيغ، وتتصل بعمليات إدارة التسجيلات الكثيرة أو مجموعات النتائج الكبيرة.

وفضلاً عن خدماته الأساسية، تتوافر للبروتوكول Z39.50 إمكانيات تصفح الكشافات، وإمكانيات ضبط عمليات الوصول Access، وإمكانيات إدارة الموارد، كما أنه يدعم الخدمات الموسعة التي تسمح بمزيد من التوسعات extensions. إنه حقًا لمعيار ضخم ومرن.

المكتبة المرجعية الفنية الشبكية لعلوم الحاسب NCSTRL، وداينيست Dienst :

"الفهرس الموحد" هو فهرس واحد يشتمل على تسجيلات مقتنيات عدة مكتبات، وتساعد هذه الفهارس الموحدة (التي تستخدمها المكتبات منذ وقت طويل قبل ظهور الحاسبات الآلية) على إجراء عملية البحث الموزع اللامركزي، من خلال تجميع المعلومات التي سيجري بحثها في فهرس واحد. وتعد خدمة البحث في الويب بمثابة فهارس موحدة للويب، حتى وإن تضمنت إحداها تسجيلات فهارس غير معدة بشكل صحيح. وثمة طريقة بديلة للبحث الموزع تتمثل في أن يكون لكل مجموعة مقتنيات كشافها القابل للبحث؛ بحيث يقوم برنامج البحث بإرسال الاستفسارات إلى هذه الكشافات المنفصلة، ثم دمج النتائج جميعها وتقديمها للمستفيد.

والمكتبة المرجعية الفنية الشبكية لعلوم الحاسب (NCSTRL) هي اتحاد لمجموعات المكتبات الرقمية الهامة للباحثين في مجال علوم الحاسب، وهي تستخدم بروتوكولاً يسمى داينست Dienst، وفي سبيل تخفيض تكاليف تقبله،

بني هذا البروتوكول على مجموعة متنوعة من المعايير الفنية المألوفة لعلماء الحاسب الذين يعتبرون من أكثر مستخدمي نظام اليونكس UNIX، والإنترنت، والويب. وتجدر الإشارة إلى أن الإصدارة الأولى من بروتوكول داينست كانت تقوم على إرسال الاستفسارات البحثية إلى جميع الحاسبات الخادمة، ونظراً لأن أعداد هذه الحاسبات تزايدت، فقد توقف العمل بهذا الأسلوب، وإذا لم يكن هناك خادم معين موجود، فإن النظام برمته يضعف. والآن يتوافر لداينست كشاف أصلي master index يعد شكلاً من أشكال الفهارس الموحدة.

اللوحة (١١-٤)

المكتبة المرجعية الفنية الشبكية لعلوم الحاسب، ونموذج داينست للبحث الموزع

NCSTRL and the Dienst Model of Distributed Searching

المكتبة المرجعية الفنية الشبكية لعلوم الحاسب هي مكتبة موزعة أو لامركزية للإنتاج الفكري في علوم الحاسب، وخاصة التقارير الفنية منها. وتقوم الجهات المتعاونة بتحميل مجموعاتها على حاسبات خادمة محلية، ويتم الاتصال أو الوصول إلى هذه الحاسبات الخادمة، إما عن طريق بروتوكول نقل الملفات FTP، أو عن طريق بروتوكول مسمى داينست Dienst، وهذا البروتوكول الأخير هو بروتوكول للمكتبة اللامركزية، وقد قام بتطويره كل من جيم ديفس Jim Davis من شركة زيروكس، وكارل لاجوز Carl lagoze من جامعة كورنيل، وتم تطوير هذا البروتوكول كجزء من مشروع التقارير الفنية لعلوم الحاسب CSTR project ، الذي أشير إليه في

الفصل الرابع. وفي بداية المشروع كانت هناك خمسة جامعات متعاونة، ازدادت لتصل إلى أكثر من مئة جهة من جميع أنحاء العالم بحلول عام ١٩٩٨م، كانت ثلاث وأربعون جامعة منها تشغل خادمات داينست. وتقوم المكتبة المرجعية الفنية الشبكية لعلوم الحاسب وداينست على الجمع بين الخدمات اليومية مع قاعدة اختبار test bed لأغراض إدارة المعلومات الموزعة. وتعد هذه المكتبة من الأمثلة القليلة لمجموعات البحث التي تقوم بتشغيل إحدى الخدمات العملية للمكتبات الرقمية.

وتقوم البنية الفنية لداينست على تقسيم خدمات المكتبة الرقمية إلى أربع فئات رئيسية هي: المستودعات والكشافات،والمجموعات، وواجهات المستفيدين. وتوفر بروتوكولاً مجانياً يعرف بهذه الخدمات؛ ويعمل هذا البروتوكول على دعم عمليات البحث الموزع للمجموعات التي تدار على نحو مستقل. ويتوافر لكل خادم كشاف للمواد التي تختزن فيه. ولكي تتم عملية البحث من خلال استخدام الإصدارات الأولية من داينست، كان برنامج واجهة المستفيدين يقوم بإرسال الاستفسار إلى جميع مواقع داينست بحثاً عن الكائنات objects التي تضاهي هذه التساؤلات، وكان على واجهات المستفيدين أن تنتظر حال وصول استجابات من الحاسبات الخادمة. وقد كان هذا هو البحث اللامركزي في شكله الأساسي، وقد بدت كثير من المعوقات بسبب الزيادة الكبيرة في أعداد الحاسبات الخادمة، وقد كانت المشكلة الأساسية هي أن جودة الخدمة التي يراها أحد المستفيدين كانت تعتمد على مستوى الخدمة التي كان يقدمها أسوأ مواقع داينست. ولفترة معينة ظل الخادم الموجود في جامعة

كارنيجي ميلون خارج حدود التعويل عليه، وإذا فشل في الاستجابة لتساؤل معين، كان على واجهة المستفيد الانتظار حتى تصدر الإشارة الآلية على انتهاء الوقت، بل إنه حتى عندما كانت جميع الحاسبات الخادمة في طور العمل، كانت تحدث في الغالب تأخيرات مطولة بسبب صعوبة الاتصال ببعض الحاسبات الخادمة القليلة.

ولا شك أن البحث البطيء يدعو إلى تذمر المستفيد، كما أن البحث يكون أكثر خطراً عندما يفقد بعض المجموعات؛ فالفشل في بحث جميع الكشافات يعني أن الباحث قد يفقد معلومات هامة، وما المشكلة ببساطة إلا بسبب الحاسب الخادم.

وقد تم إعادة تصميم داينست لتفادي هذه المشكلات، وتقسم الآن المكتبة المرجعية الفنية لعلوم الحاسب إلى مناطق، وقد كان هناك في البداية مركزان إقليميان في الولايات المتحدة، وأربعة في أوروبا. ومع هذا النموذج الإقليمي، يتم وضع الكشاف الأصلي في مكان مركزي (جامعة كورنيل)، كما يتم اختزان نسخ قابلة للبحث من هذا الكشاف في المراكز الإقليمية، بحيث يكون هناك في كل موقع إقليمي كل شيء يريد المستفيد البحث عنه، وأي معلومات يريدها. ولا يتصل المستفيد بالمواقع الفردية للمجموعات إلا من أجل استرجاع المواد المختزنة بها. وتترك مسألة اختيار أي المواقع الإقليمية التي يتم الاتصال أو البحث فيها بها للمستفيد، حيث يمكن للمستفيد، نتيجة لمشكلات الاتصال بالإنترنت، أن يحصل على أفضل الأداء من أحد المواقع الإقليمية البعيدة من تلك المواقع القريبة له من الناحية الجغرافية.

اتجاهات البحث في السبل البديلة للبحث الموزع :

إن النجاح أو الفشل الذي يحدث في سياق أحد التجمعات الاتحادية federationيعد تحدياً تنظيمياً أكثر منه تحدياً فنياً. ومن المؤكد أن بعض أعضاء هذا الاتحاد يقدم خدمات تفوق ما يقدمه الآخرون، كما أن مستويات الدعم التي يقدمها كل منهم تختلف اختلافاً كبيراً، وبالرغم من ذلك، لا ينبغي ربط جودة الخدمة بأسوأ المنظمات أداءً لوظائفها. وتصور اللوحة رقم (١١-٤) كيف أعيد تصميم نظام داينست، الذي تستخدمه المكتبة المرجعية الفنية الشبكية لعلوم الحاسب، لكي يلبي تلك الحاجة.

وتضع كل خدمة من خدمات المعلومات بعض الافتراضات الضمنية عن السيناريوهات التي تدعمها، والاستفسارات التي تقبلها، وأنواع الإجابات التي توفرها. وهذه تنفذ على أنها من قبيل التيسير، كما هو الحال بالنسبة لخدمات البحث عن المعلومات وغيرها من خدمات التصفح والتنقية والاشتقاق extract؛ فالمستفيد عادة ما يطمح إلى الحصول على معلومات متماسكة تلبي احتياجاته الشخصية، مع أن مصادر المعلومات في هذا العالم مترامي الأطراف والتي تعد متماسكة - تتفاوت فيما بينها، فكيف تستطيع إذن مجموعة متنوعة من المنظمات أن توفر للمستفيدين الحاسبات الخادمة التي تم تصميمها وفقاً لسيناريوهات ضمنية مختلفة أن تقدم استكشافاً فعالاً لمصادرها بدون عملية توحيد صارمة؟ كما أن هناك تساؤلات أو عقبات فنية صعبة تواجه إحدى الخدمات الفردية التي تدار على نحو مركزي، وقد باتت هذه القضايا معقدة فعلاً عندما خضعت مصادر المعلومات لسيطرة جهات لها استقلاليتها.

إن التحديات التنظيمية تعد بالغة الصعوبة لدرجة أنها تحد من الخيارات الفنية المتاحة، وباستثناء ما يحدث في الاتحادات القوية، فإن الأمل الوحيد في التقدم يكمن في إيجاد أُطر عمل فنية، بالتوازي مع استراتيجيات تقبل هذه الأطر، يمكن للمنظمات أن تقبلها بصورة متدرجة. وبالنسبة لكل طريقة من هذه الطرق يجب أن يكون هناك بديل ذو مستوى أدنى (عادة ما يكون هو الحالة الموجودة حالياً status quo)، وذلك حتى لا تعاق الخدمات عن أداء أي شيء جديد بالاهتمام بسبب قلة من النظم. ولذلك، فإن هناك في المكتبة المرجعية الفنية الشبكية لعلوم الحاسب - على الرغم من أن داينست يعد البروتوكول المفضل - أكثر من نصف المواقع تحمل مجموعاتها على حاسبات خادمة يدعمها بروتوكول نقل الملفات ftp.

إن كثيراً من البحوث في مجال البحث الموزع تبدأ ببناء فهارس موحدة مما وراء البيانات التي يقدمها المنشئ أو الناشر. وقد كان هذا أحد الدوافع الكامنة وراء معيار دبلن كور؛ حيث تحتاج النظم الآلية إلى جمع ما وراء البيانات هذه وضمها في كشاف قابل للبحث.

وهناك مجال آخر من مجالات البحث في هذا المجال يهدف إلى تطوير أساليب تعمل على تقييد عمليات البحث على أكثر المجموعات نجاحاً، فالمستفيدون نادراً ما يرغبون في البحث في كل مصدر من مصادر المعلومات المتاحة على الإنترنت، ولكنهم يرغبون في البحث في فئات محددة، مثل فهارس المنفردات، أو كشافات الإنتاج الفكري الطبي. ولذلك فإن هناك حاجة إلى بعض الوسائل التي تمكن المجموعات من تقديم ملخصات

لمحتوياتها. مع ملاحظة أن هذا يكتسب أهمية خاصة عندما يكون الوصول مقيداً بضابطي المصادقة وآليات دفع الرسوم. وإذا تم توفير الوصول المجاني إلى أحد المصادر، فإن برنامجاً خارجياً يمكن - على الأقل من الناحية النظرية - أن يقوم بعمل ملف إحصائي بأنواع المواد وبالمصطلحات المستخدمة. وعندما يتاح لمستفيد خارجي إمكانية الوصول فقط من خلال إحدى واجهات البحث، فإن مثل هذا التحليل لا يكون ممكناً.

وفي عام ١٩٩٦م، قام لويس جرافانو Luis Gravano بجامعة ستانفورد- بدراسة الكيفية التي يمكن بها أن يقوم أحد الحاسبات العميلة بدمج النتائج المتحصل عليها من خدمات بحث مستقلة، وقام هذا الباحث بتطوير بروتوكول مخصص لهذا الغرض عرف باسم ستارتز STARTS. وكان هذا البروتوكول مشروعًا مشتركًا بين جامعة ستانفورد من ناحية، وعدد من شركات الإنترنت الرائدة من ناحية أخرى. ويدل مدى الترحيب الذي أبدته تلك الشركات للانضمام إلى جهود هذا المشروع على أنها تدرك أن أفكاره تكتسي أهمية خاصة وأساسية بالنسبة للبحث الموسع عبر الإنترنت، وبقليل من جهود التوحيد يمكن أن تؤدي إلى تطور كبير في عمليات البحث.

وقد كان جرافانو في تحليله ينظر إلى المعلومات المتاحة عبر الإنترنت على أنها مجموعة كبيرة من المواد، وكل منها منظم على نحو مختلف، كما أن لكل منها محرك بحث خاصًّا بها. والفكرة الأساسية تتمثل في تمكين الحاسبات العميلة من استكشاف الخصائص العامة لمحركات البحث وللمجموعات التي تحتفظ بها. غير أن التحدي يكمن في اختلاف محركات

البحث، كما أن خصائص هذه المجموعات هي الأخرى مختلفة. والصعوبة ببساطة ليست في أن لواجهات التعامل أساليب مختلفة في صياغة استراتيجيات البحث فيها، بل في أن صياغة الاستراتيجية الواحدة يجب أن تعاد بعدة طرق لتتلاءم مع النظم المختلفة، كذلك فإن الخوارزميات المستخدمة هي الأخرى مختلفة تماماً، فبعض النظم تستخدم الطرق البولينية أو المنطقية في البحث، وبعضها الآخر له أساليبه الأخرى في ترتيب النتائج. كما أن محركات البحث التي تقوم بتقديم قائمة مرتبة طبقياً بالنتائج تعطى مؤشرات قليلة عن الكيفية التي تم بها ترتيب هذه النتائج. وغالباً ما يكون حساب عملية الترتيب الطبقي سراً تجارياً. ونتيجة لذلك يكون من غير الممكن دمج قوائم النتائج الطبقية التي ترد من مصادر متعددة في قائمة واحدة وبترتيب طبقي مقبول؛ ذلك لأن الترتيب الطبقي يتأثر تأثراً كبيراً بطبيعة الكلمات المستخدمة في المجموعة، ولذلك فإن دمج النتائج المتحصلة من مصدرين يستخدمان خوارزمية الترتيب الطبقي نفسها سيكون محفوفاً بالمصاعب. ويعمل بروتوكول ستارتز على تمكين محركات البحث من عمل تقارير عن خصائص مجموعاتها، وأساليب الترتيب التي تتبعها، حتى يستطيع برنامج أحد الحاسبات العميلة أن يدمج النتائج المتحصل عليها من مصادر متعددة.

اللوحة رقم ١١-٥

البنية الفنية لمشروع هارفست **The Harvest Architecture**

كان هارفست مشروعاً بحثياً عن أساليب البحث الموزع، تبناه مايكل شوارتز Michael Schwartz، ثم انتقلت رعايته إلى جامعة كلورادو، وعلى الرغم من أن هذا المشروع انتهى عام ١٩٩٦م، إلا أن الأفكار الأساسية التي

ارتبطت بتطوير بنيته الفنية لا تزال مناسبة لهذا المجال. وتركز الفكرة الأساسية لهذا المشروع على تقسيم الوظائف الأساسية في نظام البحث المركزي إلى عدة أنظمة فرعية منفصلة بعضها عن بعض، بعد تحديده بالطبع للنماذج والبروتوكولات التي تحكم عملية الاتصال بين هذه النظم الفرعية، وقد تم تطوير برنامج لعرض هذه النماذج والبروتوكولات ولكيفية تطبيقها.

ولعل أهم ما يميز مشروع هارفست هو برنامج "الجامع Gatherer"، الذي يتولى عملية تجميع المعلومات الكشفية من مجموعات المكتبات الرقمية، مع ملاحظة أن هذا البرنامج أكثر ما تكون كفاءته عندما يتم تركيبه وتهيئته مع نظام إدارة تلك المجموعات. ويقوم كل "جامع" باشتقاق معلومات التكشيف من المجموعات وإرسالها في صيغة معيارية عن طريق بروتوكول قياسي إلى برامج تسمى "بالوسطاء أو البروكر brokers"؛ حيث يقوم الوسيط ببناء كشاف تجميعي combined index يضم المعلومات المأخوذة من كل المجموعات التي تم تكشيفها.

وتجدر الإشارة إلى أن البنية الفنية لنظام هارفست تتمتع بكفاءة عند استخدامها مع شبكة من الموارد تزيد على كفاءتها عند استخدامها مع طرق التكشيف التي تعتمد على زواحف الويب، وبالرغم من أن الفريق قد عمل على تطوير الذاكرات الفورية caches، وطرق الاستنساخ من أجل رفع كفاءة هذه النظم، فإن الفائدة الحقيقية تمثلت في البحث الجيد عن المعلومات واستكشافها. ومع أن جميع برامج "الجوامع" تقوم بنقل المعلومات في بروتوكول محدد،

يعرف بالصيغة المختصرة لتبادل الكائنات the summary Object Interchange Format (SOIF)، فإن الكيفية التي تجمع بها المعلومات يمكن تفصيلها على حسب المجموعات الفردية. وفي الوقت الذي تعمل فيه زواحف الويب على المعلومات المتاحة مجاناً open access، فإن الجوامع يمكن أن تتمتع بمزايا وصول أكثر حيث يمكن استخدامها في تكشيف المعلومات المحظورة. كما أنه يمكن تهيئتها لتناسب قواعد بيانات محددة، ولا يتطلب الأمر أن يتم تقييدها على المعلومات المتاحة على صفحات الويب أو أي صيغة محددة فقط. كذلك يمكن تضمينها قواميس أو معاجم مصطلحات لأغراض البحث في مجالات موضوعية متخصصة، ولا ريب أن جمع هذه الخصائص مجتمعة تعد مزايا كبرى لهذا النظام.

ومما تجدر الإشارة إليه أن كثيرًا من مزايا البنية الفنية لنظام هارفست تتوارى ما لم يتم تركيب "الجامع" مع مجموعات المكتبة الرقمية، ولهذا السبب فإن البنية الفنية لهارفست تعد فعالة بشكل خاص بالنسبة للمكتبات الرقمية الاتحادية. ففي سياق ذلك الاتحاد يمكن أن تقوم أي مكتبة بتشغيل الجامع الخاص بها، والذي يقوم بنقل المعلومات الكشفية إلى "الوسطاء (البروكر)" الذين يقومون بدورهم ببناء الكشافات الجامعة للمكتبة كلها، ومن ثم يتم الجمع بين مزايا عملية التكشيف المحلية من ناحية، ومزايا الكشاف المركزي من ناحية أخرى.

ما وراء البحث :

إن عملية استكشاف المعلومات تعد أكثر عمقاً من عملية البحث عن المعلومات، ولذلك فإن معظم المستفيدين يستخدمون بعض أشكال الدمج بين التصفح والبحث المنظم. وقد أشرنا في الفصل العاشر إلى المتطلبات التي ينبغي أن تتوافر للمستفيدين لكي يتمكنوا من إجراء عملية البحث عن المعلومات، وإلى صعوبة تقويم فاعلية عملية استرجاع المعلومات في إحدى جلسات البحث التفاعلية مع تواجد المستفيد في الحلقة [١] with user in the loop، ولا شك أن المكتبات الرقمية اللامركزية تشعر بحدة كل تلك المشكلات.

لقد ظل التصفح على الدوام هو الطريقة المهمة لاستكشاف المعلومات في المكتبات، ويمكن أن يكون التصفح بالبساطة التي تكون عليها عملية تصفح أرفف المكتبة للتعرف إلى الكتب التي تُجمع بعضها مع بعض. غير أن هناك وسيلة أخرى أكثر منهجية تتمثل في البدء بتصفح أحد الكتب ثم الانتقال إلى الأعمال الأخرى التي يحيل إليها هذا العمل. ومن المعروف أن معظم مقالات الدوريات وغيرها من الأعمال العلمية الأخرى تشتمل على قوائم بالإرجاعات البيليوجرافية التي تحيل إلى أعمال أخرى. ومع أن تتبع هذه الإرجاعات يعد جزءاً أساسياً من عملية البحث العلمي؛ فإنها بلا شك مهمة مرهقة، وخاصة عندما تكون المواد كائنات مادية يجب استرجاعها مكتملة في آن واحد. ولعل وجود الروابط الفائقة، يجعل تتبع الإرجاعات البيليوجرافية عملية سهلة، وهناك مقولة عامة ترى أن تتبع الروابط والإرجاعات تعد من الأمور الميسورة في سياق المكتبات الرقمية، مع أن كفاءة الفهارس

(١) لعل المؤلف يقصد دائرة التفاعل بين المستفيد ونظام البحث (المترجمان).

والكشافات عادة ما تكون أكبر في سيتاق المكتبات التقليدية. ومن ثم فمن المحتمل أن يكون التصفح أكثر أهمية - بشكل نسبي - في سياق المكتبات الرقمية.

وإذا تتبع المستفيدون توليفة تجمع بين التصفح والبحث، مستخدمين أنواعاً متفرقة من المصادر ومحركات البحث، فما هي إذن درجة الثقة التي سيولونها للنتائج التي يتحصلون عليها؟ لقد أشرنا في هذا الفصل إلى صعوبات المقارنة بين النتائج التي يتم الحصول عليها من خلال عمليات البحث في مجموعات مختلفة من المعلومات، وصعوبات تقرير درجة تطابق المعلومات وتكرارها في عملين تم العثور عليهما من مصادر مختلفة. ويواجه المستفيدون الجادون من المكتبات الرقمية مشكلة بقدر ما هي دقيقة فإنها تعد أكثر خطورة، وهي أنه من الصعوبة غالباً أن نعرف مدى شمولية البحث الذي يتم تنفيذه؛ فالمستفيد الذي يقوم بالبحث في أحد مراصد البيانات المركزية، كنظام الميدلاين Medline الذي ترعاه المكتبة الطبية، يمكن أن يكون على درجة كبيرة من الثقة في أن البحث أجري لكل تسجيلة من تسجيلات النظام. قارن ذلك مع بحث موزع يجري في عدد كبير من مجموعات البيانات، وتساءل: ما فرصة ضياع معلومات هامة؟ هل لأن إحدى مجموعة البيانات تختفي في مجموعات أخرى عند تقديم المعلومات الكشفية، أم لأن إحدى مجموعات البيانات تخفق في الرد على التساؤلات البحثية ؟

إن عملية البحث الموزع تلخص الوضع الحالي للمكتبات الرقمية، فمن إحدى وجهات النظر، يمكن القول بأن أي أسلوب فني لا يخلو من نقاط ضعف خطيرة، وخاصة أنه لم تظهر بعد المعايير الفنية الكاملة المتفق عليها، كما أن فهم حاجات المستفيدين لا يزال في مراحله التمهيدية، يضاف إلى ذلك أن الصعوبات التنظيمية لا تزال متفاقمة. وعلى الرغم من ذلك، وفي الوقت نفسه، هناك كميات هائلة من المعلومات متاحة على الإنترنت، كما أن برامج بحث الويب متاحة مجاناً، بل إن الاتحادات والخدمات التجارية تتسع بشكل متزايد. وأخيراً يمكن القول إنه عن طريق الجمع الذكي بين عمليتي البحث والتصفح، يستطيع المستفيدون الشغوفون أن يصلوا إلى المعلومات التي يبحثون عنها.

لسماع الصوت، يقوم المستودع بنقل سلسلة متتابعة من المسارات الصوتية، وعندما تصل إلى حاسب المستفيد، تحول إلى صوت ثم يتم سماعها، وهذا يتم تنفيذه في مُدد زمنية قليلة، ولكن لا يوجد هناك مدقق للأخطاء، وإذا لم تصل حزمة ما عندما يحين وقت سماعها فإنه يتم تجاهلها ويسمع المستفيد انقطاعاً قصيراً أو فجوة في الصوت.

ومع أن هذه العملية تبدو غير دقيقة، فإن برنامج ريال أوديو، وخاصة عندما يكون الاتصال عبر الشبكة معقولاً إلى حد ما، يقوم بنقل الصوت بشكل مقبول جداً عبر خطوط اتصال تصل سرعتها إلى ٢٨,٨ كيلوبايت في الثانية. وقد استخدم برنامج ريال أوديو في التجارب المبكرة لتخزين فقرات صوتية وبثها عبر البرامج العامة للإذاعة الوطنية.

ويستخدم برنامج ريال أوديو أساليب الويب القياسية المعروفة في هذا الصدد، عدا خاصتين من الخصائص المطلوبة لنقل الإشارات الصوتية عبر الإنترنت في الوقت الفعلي لبثها، أولى هاتين الخاصتين أن متصفح المستفيد يتوجب عليه قبول سيل من البيانات الصوتية المحملة بصيغة ريال أوديو، وهذا يتطلب بالطبع أن يضيف المستفيد "مسجلاً player" خاصاً إلى متصفحة (ملحوظة: سجلات ريال أوديو متوافرة للتحميل عبر الإنترنت). أما الخاصية الثانية فتتمثل في وجوب استخدام المكتبة - لكي تحقق تدفقاً

متواصلاً للبيانات - بروتوكول UDP [1] بدلاً من استخدام بروتوكول الإرسال المعروف بـ TCP ، يلاحظ في هذا الصدد أن بعض نظم أمن الشبكات لا تقبل التعامل مع حزم البيانات المرسلة عن طريق هذا البروتوكول (أي UDP)، ومن ثم فإن برنامج ريال أوديو لا يمكن تداوله على نحو شائع.

الكائنات النشطة أو الحركية dynamic، والمركبة :

إن كثيراً من الكائنات الرقمية التي تشكل الآن مجموعات المكتبات الرقمية، لا يمكن تمثيلها أو عرضها على أنها ملفات بيانات ثابتة دائماً، مع الأخذ في الاعتبار ما يلي :

– **الكائنات النشطة أو الحركية dynamic:** تشمل هذه الكائنات البرامج الآلية، وتطبيقات جافا أبيليت والمحاكيات، والبيانات التي يتم الحصول عليها عن طريق أجهزة الاستشعار العلمية، وألعاب الفيديو. وعند التعامل مع مثل هذه الأنواع من الكائنات، فإن ما يعرض للمستفيد يعتمد على كفاءة أداء برامج الحاسب أو الأنشطة الأخرى الخارجية، ولذلك فإن ما يحصل عليه المستفيد من نتائج قد يختلف في كل مرة يتم فيها التعامل مع أي من هذه الكائنات.

– **الكائنات المركبة Complex Object:** يمكن لكائن مكتبي واحد أن يتكون من عدة عناصر مترابطة، ويمكن لهذه العناصر أن تكون بينها علاقات متنوعة، كما يمكن أن تكون هذه العناصر مترابطة ترابطاً عضوياً، أي مكملة

---

(١) بروتوكول إنترنت ينقل حزم البيانات دون تدقيق الأخطاء (المترجمان).

بعضها لبعض من حيث المحتوى، كما هو الحال بالنسبة لمسارات الصوت والصورة في التسجيلات المرئية، كما يمكن أن تكون عروضاً بديلة، كما هو الحال بالنسبة للصور المأخوذة عن طريق الأقمار الصناعية، فمنها ما هو عالي الوضوح، ومنها ما هو منخفض الوضوح؛ كما يمكن لهذه الكائنات أن تكون عبارة عن بدائل surrogates كما هو الحال بالنسبة لكل من البيانات، وما وراء البيانات. وكثيراً ما تكون التفرقة بين هذه البدائل صعبة وخاصة عند التعامل معها عن قرب، ومثال ذلك التساؤل: هل تعد الصورة المنمنمة عرضًا بديلاً ؟ أم تعد ما وراء بيانات لصورة أكبر ؟

- **أساليب البث التبادلية Alternate dissemination**: يمكن للكائنات الرقمية أن تقدم للمستفيد خيارات سبل إتاحتها؛ حيث يمكن لأحد الكائنات المكتبية - على سبيل المثال - أن تقدم حالة الطقس في مطار سان فرانسيسكو، وعندما يصل المستفيد إلى هذا الكائن، فيجد المعلومات المسترجعة في شكل بيانات (كدرجة الحرارة، ودرجة الرطوبة، وسرعة الرياح واتجاهها)، أو في شكل صورة فوتوجرافية تظهر شكل السحب. مع ملاحظة أن المعلومات تتم قراءتها عن طريق أجهزة الاستشعار (عند طلب ذلك) أو من خلال الجداول التي يتم تحديثها بانتظام.

- **قواعد البيانات Databases**: تضم قاعدة البيانات العديد من التسجيلات البديلة، بالإضافة إلى التسجيلات الفردية المختلفة التي يتم اختيارها في كل مرة يتم الوصول فيها إلى هذه القاعدة. وهناك بعض قواعد البيانات يمكن اعتبارها نموذجاً حقيقياً لمجموعات مكتبة رقمية متكاملة، مع التسجيلات

الفردية التي تمثل الكائن الرقمي داخل هذه المجموعات، وفي المقابل، هناك قواعد بيانات أخرى - كالأدلة مثلاً - يمكن أن تعد كائنات مكتبية قائمة بذاتها أو كما هي.

وتجدر الإشارة إلى أن أساليب إدارة مثل هذه الكائنات شديدة العمومية more general object لا تزال محل جدل بين أوساط المهتمين؛ ففي الوقت الذي تتيح فيه الويب إطار عمل موحد يستخدمه معظم الناس عند تعاملهم مع الملفات الثابتة، لا نجد مثل هذا الإطار المقبول على نطاق واسع بالنسبة لهذه الكائنات العامة. وبصرف النظر عما يحدثه استعمال المصطلحات من تفاوت في دلالات المعاني، فإن هناك مجموعة من وجهات النظر تتفق على ربط البعد الفكري للمواد المكتبية بالبناء الداخلي لها، وهو ما يعرف في كثير من الأحيان "بالنموذج الوثائقي document model"، ولكن لأن هذا المصطلح قد ينطبق على جميع جوانب المكتبات الرقمية، فإن هناك مصطلحاً آخر نرى أنه أفضل منه وهو "النموذج الكائني object model".

تحديد الهدية أو التعريف Identification :

تتمثل أولى خطوات بناء نموذج كائني في إيجاد طريقة للتعريف بهذه المواد أو بتحديد هويتها، ويستخدم "المُعرف identifier" للإشارة إلى الكائنات سواء في الفهارس، أو في الإشارات الببليوجرافية، وفي تخزين هذه الكائنات وإتاحة الوصول إليها، كما تستخدم في عملية إدارة الوصول إلى هذه الكائنات، وفي أرشفتها أو حفظها لفترات طويلة. ومع أن هذا قد يبدو بسيطاً، فإن المُعرفات يجب أن تكون مستوفية للمتطلبات التي تجعلها قادرة

على تحديد هوية المواد التي تعبر عنها. وثمة بعض الموضوعات في مجال المكتبات الرقمية تثير كثيراً من المناقشات الساخنة، ومن هذه الموضوعات قضيتا الأسماء names والمعرفات identifiers. مع ملاحظة أنه لم يظهر حتى الآن توجه دولي يتصل بعملية التسمية بالشكل الذي يلبي كل الاحتياجات.

ويتمثل أحد جوانب الجدل المثار حول هذا الموضوع في التساؤل عما إذا كان من الضروري تضمين الاسم ألفاظًا دلالية semantics أم لا؟ فهناك من يؤيدون ويدافعون عن فكرة الأسماء كاملة الدلالة completely semantic names، وخير مثال على ذلك "مُعرف العناوين والمساهمات المسلسلة the serial item and contribution identifier ((SICI))، الذي يتبنى تحديد هوية أحد الأعداد المنشورة من دورية معينة، أو إحدى المقالات المنشورة فيها، اعتماداً على مجموعة محددة ودقيقة من القواعد. ومع أن العادة قد جرت على إمكانية اشتقاق هذا المحددٌ (SICI) من مقالة الدورية أو من إحدى الإشارات المرجعية، فإن هذا الهدف قد يصعب تحقيقه في بعض الأحيان، وإذا كان هذا المعيار قد حقق نجاحاً في هذا الصدد، فإن نجاحه هذا تحقق بفضل وجود معيار آخر يتم بناء عليه تحديد هوية عناوين المسلسلات على نحو متفرد. والمثال التالي عبارة عن نموذج فعلي مُعرف العناوين والمساهمات المسلسلة (SICI)، في تعريفه لهوية مقالة منشورة في إحدى الدوريات التي ينشرها جون ويلي John Willy and Sons.

0002-8231(199601)47:1<23:TDOMII>2.0.TX;2-2

إن الأسماء كاملة الدلالة - كما هو الحال بالنسبة لمحدد العناوين والمساهمات المسلسلة - يقتصر تطبيقها بشكل تام على الفئات الضيقة من

المعلومات كما أنها تبدو طويلة وغير مستساغة بسبب تعقد القواعد التي تستخدم في إنشائها. ونتيجة للصعوبات التي تكتنف استحداث المحددات الدلالية الخاصة بالكائنات شديدة العمومية، والتي تزيد المنازعات على العلامات التجارية وعلى الأسماء الأخرى من تعقيدها، فإن هناك من يؤيدون البديل الآخر المعاكس لهذا الأسلوب، المتمثل في إيجاد محددات عشوائية random identifiers لا تتضمن بالضرورة معلومات دلالية تشير إلى من وضع هذه الأسماء، ولا إلى ما تشير. ويمكن بذلك أن تكون التتابعات المتتالية العشوائية string التي تستخدم كمحدِّدات أقصر من النوع الأول، إلا أنه يصعب على الناس تذكرها، بل قد يصعب على الحاسبات معالجتها بسبب عدم اشتمالها على المعلومات اللازمة لذلك.

وإذا كان هناك في الواقع الفعلي كثير من الأسماء الرمزية mnemonic أي التي تشتمل على المعلومات التي تكفل سهولة تذكرها، مثال ذلك اسم النطاقwww.apple.com، حيث يبدو للوهلة الأولى أن هذا الاسم هو اسمٌ دلالي، أي اسم يشير إلى موقع على الويب لمؤسسة تجارية تسمى apple، بالرغم من ذلك يمكن القول إن هذا تخمين يساوره شيءٌ من المعرفة، حيث تستخدم البادئة www أساساً للإشارة عادة إلى مواقع الويب، مع أن هذا مجرد اتفاق، ولكن لأن هناك العديد من المؤسسات التجارية الموسومة بآبل apple، فإن هذا الإسم السابق لا يعطي أي إشارة صريحة إلى ما إذا كان هذا الموقع تديره شركة Apple Computer أو غيرها من الشركات الأخرى.

وثمة مشكلة أخرى تتمثل في صعوبة الجزم بما يشير إليه الاسم، بمعنى أن الاسم قد لا يتضح منه هل هو يشير إلى عمل، أم إلى أحد أشكال التعبير،

أم إلى بيان الحالة، أم إلى عنصر معين. قارن ذلك مثلاً بما يحدث في سياق "الأرقام المعيارية الدولية للكتب International standard book number (ISBN)" التي طورها ناشرو الكتب وموزعوها لاستخداماتهم الخاصة، حيث تهدف هذه الأرقام إلى التمييز بين منتجات هؤلاء الناشرين، التي تكون محل تداول بالبيع أو الشراء؛ حيث تأخذ - على سبيل المثال - إحدى طبعات الكتاب التي صدرت مغلفة تغليفاً مقوىً رقماً معيارياً (ISBN) يختلف عن الرقم الذي تأخذه طبعة أخرى من هذا الكتاب تكون قد صدرت مغلفة ولكن تغليفاً ورقياً، مع أن هاتين الطبعتين متطابقتان في المحتوى العلمي. ومع ذلك، فقد تجد المكتبات في هذا التمييز ما لا طائل من ورائه، ذلك لأن التمييز الطبيعي بالنسبة للأغراض الببليوجرافية ينبغي أن يكون بين الطبعات التي تختلف في محتواها العلمي وليس في شكلها المادي، أما بالنسبة لأغراض إدارة المجموعات، فإن كل نسخة منفردة من العمل تعامل على أنها مختلفة عما سواها من النسخ الأخرى، ومن ثم تحتاج إلى محدِّد خاص بها.

أسماء النطاقات DNS والمحدِّدات الموحدة لمواقع المصادر : URLs

تعد أسماء النطاقات والمحدِّدات الموحدة لمواقع المصادر من أكثر المحددات تداولاً على الإنترنت. وبالرغم من أن المحددات الموحدة توسع مفهوم أسماء النطاقات في عدة اتجاهات، فإنهما يعدان بمثابة امتدادات أو توسعات للمفهوم الأساسي لعملية توفير اسم محدد لأحد المواقع على الإنترنت.

وقد ثبت نجاح المحدِّدات الموحدة لمواقع المصادر URLs؛ إذ تسمح

بإنشاء أي عدد من التطبيقات متعددة الاستخدام على الإنترنت، وبالرغم من ذلك، تبين أنها تسبب مشكلة طويلة الأجل للمكتبات الرقمية، وهي مشكلة مدى الثبات والاستمرارية persistence؛ ذلك أن المستفيدين من المكتبات الرقمية يطمحون دائماً أن تتاح لهم فرصة الوصول إلى مواد الإنترنت على نحو دائم وعلى فترات زمنية طويلة؛ وتعمل المحددات الموحدة على تعريف المصادر أو تحديد هويتها من خلال تحديد موقعها المستمد من أحد أسماء النطاقات، إلا أن هذا المحدد الموحد يفقد صلاحيته أو قيمته إذا اختفى اسم النطاق ولم يعد متوافراً على الإنترنت، أو إذا انتقل المصدر إلى موقع آخر مختلف. وهناك مثال شهير يعود إلى الفترات المبكرة من عمر الويب؛ فبعد تأسيس اتحاد الويب العالمي WWWC في معهد ماساشوستيس للتقنية MIT عام ١٩٩٤م، نقلت جميع جهود التوثيق المميزة من مقرها الرئيسي بالمركز الأوروبي للبحوث النووية CERN في جنيف إلى مقر هذا المعهد (MIT)، ومن ثم فقدت جميع الروابط الفائقة التي تشير إلى المركز الأوروبي للبحوث النووية صلاحيتها.

وثمة تقليد يتبع في مثل هذه الحالات، يقضي بوضع صفحة تفيد بانتقال الموقع إلى مكان آخر، غير أن العناوين التي تقوم بعملية التمرير أو التحويل forwarding addresses تختفي مع مرور الوقت أو تصبح سلاسل إجراءات مطولة. ومن ناحية أخرى، إذا حدث أن أُلغي أحد أسماء النطاقات - بسبب خروج الشركة من سوق العمل- فإن جميع المحددات الموحدة المبنية على هذا الاسم ستلغى أو تنهار هي الأخرى إلى الأبد، ولتخفيف حدة هذه المشكلة أو على الأقل هذا الوضع، ثمة عدة أشكال مختلفة من الأسماء البديلة aliases التي

يمكن استخدامها مع أسماء النطاقات ومع المحددات الموحدة، إلا أن جميع هذه الأسماء هي أقرب إلى الخدع منها إلى الحلول.

اللوحة رقم ١٢-٤

أسماء النطاقات

يتمثل الهدف الأساسي من أسماء النطاقات في تعريف هوية الحاسبات الموجودة على الإنترنت أو تحديدها عن طريق إعطائها اسماً، وليس عن طريق عنوان بروتوكول الإنترنت IP، ومن أهم مزايا هذه الطريقة أنها تكفل ثبات الاسم وعدم الحاجة إلى تغييره في حالة تغيير النظام الآلي. فعلى سبيل المثال، ظل اسم النطاق "Library.dortmouth.edu"- وعلى مدى عدة سنوات - يشير إلى سلسلة من الحاسبات بعناوين إنترنت مختلفة أو متغيرة، دون علم المستفيدين بالتغيرات التي حدثت على عناوين الإنترنت الخاصة بهذه الحاسبات.

ومع مرور الوقت، أصبحت أسماء النطاقات أكثر مرونة، ولم يعد بالضرورة أن يشير أحد أسماء النطاقات إلى حاسب معين على الإنترنت، كما يمكن لأسماء نطاقات متعددة أن تشير إلى الحاسب نفسه، كما يمكن لاسم نطاق واحد أن يشير إلى إحدى الخدمات التي تتوزع عبر مجموعة من الحاسبات.

إن تخصيص أسماء النطاقات يكوّن بناءً هرمياً، بحيث ترد في قمة هذا البناء أسماء النطاقات الجذرية root DN، ويعتمد في بناء إحدى أسماء النطاقات الجذرية على نوع المؤسسة التي تشير إليها، ومن أمثلة ذلك: الرمز .com الذي يشير إلى أن الجهة ذات طابع تجاري، والرمز .edu الذي يشير

إلـى أن الجهة ذات طابع تعليمي، والرمز .gov الذي يشير إلى أن الجهة ذات طابـع حكومـي، والرمـز .net الـذي يشير إلـى أن الجهـة تقدم خـدمات شبكيـة، والرمـز .org الـذي يشير إلى الجهات الأخرى. وهناك سلسلة ثانية من أسماء النطـاقات الجذريـة التـي تبنـى على أساس جغرافي، مثال ذلك: .ca التي تشير إلـى كـندا، .jp التــي تـشير إلــى الـيابان، .nt التــي تـشير إلــى هولـندا. ويـتم تخصيص أسماء النطاقات للمنظمات تحت واحد من هذه الأسماء الجذرية، مـثل: .Cmn.edu التـي تشير إلى جامعة كارنيجي ميلون، و.Elsevier.net التي تـشير إلـى الناشـر إلسفير، و.Loc.gov التـي تـشير إلـى مكتبـة الكونجـرس، و .Dlib.org التي تشير إلى مجلة المكتبات الرقمية.

وتجـدر الإشــارة إلــى أن أسـماء الـنطاقات فــي الـولايات ومـن الناحـية التاريخية، كـان تخصيصهـا يتـم اعتمـاداً على مبـدأ "من يتقـدم أولاً، تقدم لـــه الخدمـة أولاً "First-come, first-served". وقد كانـت إجـراءات ضـبط عملية الحصـول علـى أسماء النطـاقات - بمعنـى من يعطى اسم نطاق معين، ومـا ينبغـي أن يكـون عليه هذا الاسم - ضـعيفة إلـى حد ما، كما كانت رسوم الحصول علـى هذه الخدمـة قليلة. ومـن ثم كـان يمكـن لشخص ما من سكان مدينـة بتسبرج أن يسجل اسم النطاق "Pittsburgh.net". وفضلاً عن استشراء ظاهـرة الأسـماء غيـر المستـساغة، فقـد أدى عـدم توافـر الإجـراءات الكافيـة لضبط عملية تسجيل أسماء النطاقات إلى منازعات على العلامات التجارية، وإلى منازعات أخرى في غير ذلك.

اللوحة رقم (١٢ـ ٥)

المعلومات التي يتضمنها المحدد الموحد لموقع المصدر

إن سلسلة التمثيلات التـي تكوّن أحد المحددات الموحدة لا يتم تـركيبها عبثاً؛ وإنمـا بدقـة عالية. فالمحدد الموحد يشتمل على البيانات التـي تصف: البـروتوكول المستخدم، واسم الملـف، والخـيارات التـي تـتاح للوصـول إلـى الملف. وكل منها يشتمل على ما يلي :

– **الـبروتوكولات:** يعـد البـروتوكول أو الخدمـة هـو الجـزء الأول الـذي يـتكون مـنه المحدد الموحد، وهو ينتهي بعلامة الشارحة (:)، ومن أمثلة تلك البروتوكولات: ‎: http‏، و ‎:mailto‏، و‎:ftp‏ .

- **المحـددات المـوحدة المطلقة والنـسبية Absolute and relative URLs :**

يمكـن أن يـشير المحـدد المـوحد إلـى الملـف إمـا عـن طـريق اسـم الـنطاق الخـاص بـه، وإمـا عن طـريق تحديـد موقعه بالنسبة لملف آخر، ولذلك إذا أُتْبِعَ اسـم البـروتوكول بـشرطتين مائلتـين (//)، فـإن هـذا يعنـي أن المحـدد المـوحد يـشتمل علـى اسـم نطـاق كامل، مثال ذلك : "http://\www.dlib.org/figure.jpg"، أمـا غيـر ذلـك، فيعنـي أن العنوان يكون منـسوباً إلى الدليل الحالي، حيث يشير للبـيان الرمـزي anchor التالـي علـى سـبيل المـثال : <a href = "figure.jpg"> والموجـود فـي إحـدى صـفحات المكتوبة بلغـة html، إلـى أن ثمـة ملفـاً وهـو "figure.jpg" موجود في الدليل نفسه.

- **الملفـات :** يعمـل المحـدد الموحد على التعريف بهوية ملف معين على

أحد النظم الآلية المعروفة، حيث تعتبر المقطع "www.dlib.org" الموجود ضمن المحدد "http:\\www.dlib.org/content.html" اسم نطاق يحدد هوية أحد الحاسبات على الإنترنت، أما المقطع "content.html" فهو اسم ملف موجود على هذا الحاسب.

– **المنفذ Port :** يمكن لأحد الحاسبات الخادمة على الإنترنت أن يقدم عدة خدمات في آن واحد، حيث يقوم بروتوكول التحكم بالإرسال (TCP) بتوفير أحد المنافذ، والذي يحدد بدوره الخدمة التي يجب استخدامها. وهذا المنفذ يتم تعيينه عن طريق الشارحة (:) متبوعة برقم في آخر اسم النطاق، على سبيل المثال، المحدد "http:\\www.dlib.org:80/index.html" يتضمن إشارة إلى المنفذ المعبر عنه بالرقم 80، والذي يعد المنفذ الافتراضي لبروتوكول نقل الملف الفائق (http)، وبالتالي يمكن إلغاء هذا المنفذ من هذا المحدد على وجه الخصوص.

- **المتغيرات أو المعالم Parameters :** يمكن أن تضاف بعد المحددات الموحدة أنواع متعددة من المتغيرات أو المعالم، وذلك إما بعد الرمز "#" أو الحرف "a"، وهذه المعالم تمرر إلى الخادم عندما يتم الوصول إلى الملف.

الأسماء الدائمة أو الثابتة Persistence names، والأسماء الموحدة للمصادر URNs :

أولى مجتمع المكتبات الرقمية وناشرو المواد الإلكترونية الأسماء الدائمة اهتماماً كبيراً، وتعرف هذه الأسماء الدائمة في بعض الأحيان بالأسماء الموحدة للمصادر، وفكرة هذه الأسماء بسيطة، مفادها أن الأسماء

ينبغي أن تكون فريدة على المستوى العالمي من ناحية، وثابتة غير متغيرة مع مرور الوقت من ناحية أخرى، بمعنى أنها تكون أكثر ديمومة من أي نظام برمجيات موجود هذه الأيام، بل تكون أطول من عمر الإنترنت نفسها.

ويجب أن يكون الاسم الثابت كفيلاً بالإشارة إلى أي مصدر، أو مجموعة مصادر متاحة على الإنترنت. ومن إحدى تطبيقات الأسماء الموحدة للمصادر استخدامها للإشارة إلى المواقع الحالية لنسخ أحد الكائنات، والتي يتم تعريفها عن طريق قائمة بالمحددات الموحدة لمواقع المصادر. وهناك تطبيق آخر لهذه الأسماء الموحدة يتمثل في توفيرها لعناوين البروتوكول الإلكتروني التي لا يتطلب الأمر تغييرها عندما يقوم الشخص بتغيير الوظائف أو فردي خدمات الإنترنت؛ وهناك إمكانية أخرى للأسماء الموحدة تتمثل في توفيرها المفاتيح العامة public key لخدمات معروفة. وتصف اللوحة رقم (١٢-٦) نظام "المعالجة أو الهاندل handle"، الذي يعتبر أحد أنظمة إنشاء الأسماء الدائمة وإدارتها، كما تصف اللوحة كذلك كيفية استخدام الناشرين لهذا النظام لأغراض إدارة محددات الكائنات الرقمية.

اللوحة رقم (١٢–٦)

المعالجات أو الهاندلز ومحددات الكائنات الرقمية

المعالجات (الهاندلز) هو نظام تسمية، طوره ائتلاف مبادرة البحث الوطني Coalition National Research Initiative كجزء من نظام أو إطار عمل اقترحه كل من روبرت كهن Robert Kahn من هذا الائتلاف، وروبرت ويلنسكي Robert Wilensky من جامعة كاليفورنيا في بيركلي. وعلى الرغم من أن هذا النظام تم تطويره بعيداً عن الأفكار الأساسية التي بني عليها نظام

"الأسماء الموحدة للمصادر"، فإنه متوافق مع هذا النظام الأخير؛ بل يمكن القول إنه أول نظام للأسماء الموحدة يستخدم في المكتبات الرقمية.

ويتوافر لنظام المعالجات (الهاندلز) خطة أسماء تسمح للهيئات المستقلة باستحداث أسماء معالجة على نحو دقيق handle names بحيث تكون هذه الأسماء فريدة، كما يتوافر له نظام آلي لا مركزي يقوم باستخدام المعالجات أو الأسماء المعالجة مع البيانات التي تحيل إليها (مثل الأماكن التي يتم اختزان المواد بها)، كما تتوافر له مجموعة إجراءات إدارية تكفل بقاء دقة المعلومات على مدى فترات زمنية طويلة.

وفيما يلي مثالان فعليان للأسماء المعالجة، وهما :

hdl:cnri.dlib/magazine

hdl:loc.music/musdi.139

وواضح من المثالين السابقين أن المعالجات (الهاندلز) تتكون من ثلاثة أجزاء، يشير الأول منها إلى ما يفيد بأن هذا الصف من البيانات من نوع المواد المعالجة، أما الجزء الثاني (cnri.dlib) في المثال الأول، و loc.music في المثال الثاني فيعرف بهيئة التسمية naming authority، ويتم تحديدها على نحو هرمي، حيث يتم تحديد الجزء الأول من هيئة التسمية (أي cnri، loc) من قبل الجهة المركزية؛ أما الجزء الباقي (أي cnri.dlib) فيتم تحديده محلياً أي من قبل الجهة نفسها؛ والجزء الثالث من هذا الصف والذي يرد بعد الشرطة المائلة فهو عبارة عن مجموعة من التمثيلات التي تتفرد بها المعالجات (الهاندلز) وتعد فريدة لجهة التسمية.

ويتم ترجمة أو تفسير الاسم المعالج resolving عن طريق إرساله إلى نظام الحاسب الآلي، ثم تلقى البيانات المختزنة، حيث يوجد هناك نظام حاسوبي مركزي يعرف "بالمسجل العالمي للمعالجات the global handle registry"، ومع ذلك، فمن حق أي جهة أن تتمتع بحرية إنشاء خدمة معالجة محلية local handle service على الحاسبات الخاصة بها لمعالجة المواد، وتقديم خدمات التفسير أو المطابقة resolution. ولكي يتم ذلك يجب أولاً أن يتم تعيين هيئات التسمية العليا top-level naming authorities، وأن تكون جميع هيئات التسمية مسجلة في الخدمة المركزية، ولضمان دقة الأداء وموثقيته، يجب أن توزع كل واحدة من هذه الخدمات على حاسبات كثيرة، وأن يتم استنساخ البيانات آلياً. وقد تم في هذا الصدد توفير العديد من خدمات "الذاكرة المؤقتة أو الخفية caching"، كما هو الحال بتلك الخدمات الموصلة بالحاسبات plug-ins والتي تمكن متصفحات الويب من تفسير المعالجات أو تفسيرها.

وفي عام ١٩٩٦م، تبنت مبادرة من جمعية الناشرين الأمريكيين Association of American publishers، المعالجات وسائل لتحديد هوية المواد المنشورة إلكترونياً، وقد عرفت هذه المحدّدات "بمحددات الكائنات الرقمية DOLs"، وقد أدت هذه المبادرة إلى إنشاء مؤسسة دولية تعمل على تطوير هذه المحددات. كما تم تخصيص هيئات تسمية رقمية numeric naming للناشرين؛ حيث خصص الرقم "10.10006" - على سبيل المثال - للناشر Academic press، وبناء على ذلك فإن محدد الكائن الرقمي لأيٍّ من الكتب التي ينشرها هذا الناشر سيكون على النحو التالي

Doi:10.10006/0121585328. ويبدو جلياً أن استخدام الأرقام لهيئات التسمية تعكس الرغبة في الإقلال من المعلومات اللفظية إلى أبعد حد، ولأن الناشرين غالباً ما يقومون بعمليات إعادة تنظيم بعد الأعمال أو الجمع بينها، أو تحويلها إلى ناشرين آخرين، ولأن محددات الكائنات الرقمية هذه لا تتغير مع مثل هذه التغييرات، فإنها لا تركز على اسم الناشر الأصلي.

النظم الحاسوبية لتفسير الأسماء Computer Systems for Resolving Names :

بصرف النظر عن نظام المحددات المستخدم، يجب أن تكون هناك وسيلة سريعة وكفء يستطيع الحاسب الموجود على الإنترنت أن يستكشف ما يشير إليه اسمٌ معين؛ وتعرف هذه العملية بعملية " الترجمة أو التفسير resolving"، وتعني عملية تفسير أحد أسماء النطاقات توفير عنوان أو عناوين بروتوكول الإنترنت للنظام الآلي مع ذلك الاسم، كما تعني عملية التفسير لأحد الأسماء الموحدة للمصادر توفير البيانات المرتبطة بهذا الاسم، والتي غالباً ما تكون عبارة عن "محدد موحد URL" يقوم بتحديد الموقع الذي يُختزن فيه المصدر في الوقت الحالي.

وحيث تتوافر لكل حاسب على الإنترنت - تقريباً - البرمجية اللازمة لتفسير أسماء النطاقات، ولمعالجة المحددات الموحدة لمواقع المصادر، فقد حاولت كثير من الجماعات استخدام هذه الآليات المتوافرة في إنشاء نظم تعريف أو تحديد هويات المواد في المكتبات الرقمية. ومن إحدى هذه الآليات نظام "المحدد الموحد الدائم لمواقع المصادر persistent unified resource

locator" والمعروف اختصاراً بكلمة بيرل PURL، والذي طوره مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC.

ونظام بيرل عبارة عن محدد موحد لموقع المصدر يرد على هذا النحو التالي:

http://purl.loc.org/catalog/item1 . وفي هذا المحدد يتبين أن المقطع "purl.loc.org" هو اسم النطاق الخاص بأحد الحاسبات الذي يتوقع أن يكون اسماً دائماً، وعلى هذا الحاسب، يقوم الملف المسمى " catalog/item1" باختزان محدد موحد للموقع الذي يتم فيه اختزان العنوان item في الوقت الحالي. وإذا حدث أن انتقل العنوان إلى مكان آخر، فإن هذا المحدد يجب أن يتغير، ولكن المحدد الموحد الدائم purl - أي الاسم الخارجي- يظل دون تغيير.

ويضيف نظام المحددات الموحدة الدائمة حيلة ظريفة أو مثيرة للاهتمام بالنسبة لإدارة الأسماء، حيث تبدأ نظم التسمية الأخرى أولاً بعمل مجموعة منسقة وحيدة من الأسماء المناسبة لمجتمع كبير قد يكون هو العالم كله. ويمكن اعتبار هذا الأسلوب من أساليب البدء من أعلى إلى أسفل top-down، أما نظام المحددات الموحدة الدائمة فتعتبر من أساليب "البدء من أسفل إلى أعلى bottom-up". وبما أن كل خادم من خوادم المحددات الموحدة الدائمة يكون مستقلاً بذاته، فلا يكون هناك حاجة لتنسيق عملية تخصيص الأسماء فيما بين هذه الخوادم. ويمكن للأسماء أن تتكرر، أو أن تستخدم بدلالات لفظية مختلفة، اعتماداً على القرارات التي تُتخذ محلياً بشكل كامل. وفي مقابل ذلك، عندما يقوم أحد من الناشرين باستحداث محددات لكائنات رقمية،

فإنـه يكـون قـد شـرع بـذلك فـي بنـاء مجموعـة وحيدة من الأسماء التي يُضمن لها أن تكون متفردة عالمياً.

ما وراء البيانات البنائية ونماذج الكائنات :

أنواع البيانات data types :

أنـواع البيـانات هـي مـا وراء بيانات بنائية تستخدم لوصف مختلف أنواع الكائنـات الموجـودة فـي المكتبـة الـرقمية. وقـد نـاقشنا فـي سياق الفصل الثانـي أهمية معيار مايم MIME في تعريف نوع بيانات الملفات التي يتم تبادلها عن طـريق البـريد الإلكترونـي أو التـي تـستخدم فـي الويب. ويعد معيار مايم مثالاً جـيداً للمعاييـر التـي تتمـتع بالمـرونة الكافـية وبالـشكل الـذي يجعلـه يغطـي مجمـوعة كبيـرة من التطبيقات التي لا تزال تتسم بقدر كاف من البساطة التي تجعل من السهولة بمكان تضمينها في النظم الآلية.

يـتألف الـنموذج الكائنـي علـى الـويب مـن مجمـوعة مـن الملفـات الفائقـة والمترابطة hyperlinked من البيانات، وكل ملف من هذه الملفات ترافقه أنواع بـيانات معـتمدة فـي بنيـتها علـى معيار مـايم MIME data type، بحـيث تقـوم الـروابط الفائقـة بـتحديد طبـيعة العلاقـات بـين الملفـات، فـي حـين تقـوم أنـواع البـيانات بـإعلام واجهـة المـستفيد بالكيفـية التـي يـتم بهـا إحـضار الملـف أو استحضاره لكي يعرض للمستفيد.

الكائنات المركبة Complex Objects :

هناك كثير من مواد المكتبات الرقمية لا تكون بالبساطة التي تكون عليها الملفـات التـي يتم عرضـها عن طـريق أنـواع بـيانات المـايم البسيطة، بل هـي

أكثر تعقيداً من ذلك؛ حيث يمكن أن تتكون هذه الملفات - على سبيل المثال - من عدة عناصر بأنواع بيانات مختلفة، كما هو الحال في الصور الموجودة داخل أحد الملفات النصية، أو أن تكون هذه الملفات مرتبطة بملفات أخرى ووفق نوع معين من العلاقة، كعلاقة الجزء بالكل، أو علاقة التسلسل المنطقي. كما أن النص المرقمن يمكن أن يتكون من صفحات، وفصول، وصفحات تمهيدية، وكشاف، ومواد توضيحية، وغير ذلك من العناصر الأخرى، كما أن إحدى المقالات المنشورة في دورية إلكترونية متاحة على الخط المباشر قد تختزن في نظام آلي كعدة ملفات تشتمل على نص المقالة وما يصاحبه من صور، بالإضافة إلى مجموعة روابط معقدة تربط بين هذه الملفات. وبحكم سهولة تغيير المواد الرقمية، فمن المنطقي استحداث إصدارات مختلفة من المادة الواحدة وعلى نحو مترافق مع هذه التغيرات.

كما أن عنواناً واحداً يمكن أن يختزن في عدة صيغ رقمية تبادلية، وعندما تحول المادة الموجودة إلى صيغة رقمية، فإن العنوان نفسه في شكله المادي قد يحول عدة مرات. وخير مثال على ذلك، صورة فوتوجرافية تم التقاطها عن طريق المسح الضوئي، يمكن أن تتوافر لها عدة نسخ، منها نسخة عالية الوضوح تستخدم لأغراض الحفظ الأرشيفي، ونسخة أخرى متوسطة الوضوح، ونسخة ثالثة منمنمة. وقد تكون هذه الصيغ - في بعض الأحيان - متكافئة تماماً، كما هو الحال مثلاً بالنسبة لإحدى الصور عندما تتاح مرة مضغوطة، ومرة أخرى تختزن بشكل مضغوط ومحكم أي بدون فقدان خصائصها lossless compression، كما يمكن لهذه الصيغة كذلك أن تتحول إلى صيغة أخرى. وفي بعض الحالات الأخرى، قد تحتوي هذه

الصيغ المختلفة على معلومات مختلفة كذلك، كما هو الحال مثلاً بالنسبة لأشكال التقديم المختلفة لإحدى الصفحات النصية، حيث تعرض مرة بصيغة اللغة المعيارية الموحدة للترميز، ومرة أخرى بصيغة بوست سكريبت.

اللوحة رقم (١٢-٧)

معيار مايم **MIME**

قام كل من ناثينال بورنستين Nathaniel Borestein، ونيد فريد Ned Fred ، بتطوير مواصفة مايم، التي تعنى بامتدادات بريد الإنترنت متعدد الأغراض، وبالرغم من أن هدفها الأساسي كان منصباً على خدمة البريد الإلكتروني، فقد ثبت أنها صالحة للاستخدام مع تطبيقات أخرى كثيرة على الإنترنت، بل تعد هذه المواصفة واحدة من الأسس البسيطة والمرنة التي بنيت عليها الويب والتي أدت من ناحية أخرى إلى نجاحها.

وتجدر الإشارة إلى أن مواصفة مايم في صيغتها الكاملة تعد معقدة نظراً للحاجة إلى تواؤمها مع أنواع كثيرة من نظم البريد الإلكتروني، غير أن الأساس الذي يهم المكتبات الرقمية هو ما يتمثل في المفهوم الذي تسميه مايم "نوع المحتوى content-type"، حيث تقوم مايم بوصف نوع البيانات في ثلاثة أجزاء، كما هو الحال في المثال التالي :

Content-type:text/plain;charset = " US-ASCII "

فمن الواضح أن نوع البيانات في مواصفة مايم تتكون من: نوع البيانات أو طبيعتها (كما هو الحال بالنسبة لكلمة "نص text" في المثال)، والنوع الفرعي للبيانات (كلمة "عادي plain" التي تبين نوع النص)، ثم المعلم أو المعالم الاختيارية. فالمثال السابق يعرف نصاً عادياً كتب باستخدام مجموعة الكود

الأمريكي لتبادل المعلومات والمعروف اختصاراً بآسكي ASCII، وتشمل الأنواع شائعة الاستخدام ما يلي :

Text / plain, text/html, image/gif, image/jpeg, image/tiff, audio/basic, audio/wav, video/mpeg, video/quick draw.

وتعمل نوعية التطبيق على تقديم نوع البيانات للمعلومات التي سيتم استخدامها مع بعض البرامج التطبيقية. ويكون نوع البيانات الذي تحدده مايم بالنسبة للملفات في صيغة كل من: بي دي إف pdf، ومايكروسوفت ورد mw، وباوربوينت pp، على النحو التالي وبالترتيب نفسه.

application/pdf, application/msword, application/ppt

ولأن الملف الموجود في صيغة مايكروسوفت قد يحتوي على معلومات غير نصية، تتطلب برنامجاً آلياً خاصاً لقراءتها، فإن الإشارة إلى نوعية التطبيق هكذا " application/msword " تعني أن هذا الملف ليس ملفاً نصياً.

وعلى الرغم من أن الأمثلة السابقة قد أقرت من خلال العملية الرسمية لتسجيل أنواع المايم، فإنه من الممكن أيضاً استحداث أنواع رئيسية وفرعية غير رسمية مع أسماء الملفات، وهي تبدأ عادة بالرمز (x-) كما في المثال " audio/x-pn-realaudio".

إن النجاح الذي تحقق لمعيار مايم يعد بحق درساً لكيفية تحول أحد الأفكار الجيدة إلى نظام يتم تطبيقه وتبنيه على أوسع نطاق. ومن النقاط التي تجدر الإشارة إليها في هذا الصدد أن مواصفة مايم تجتهد اجتهاداً كبيراً لتبقى متوافقة على نحو دائم مع النظم التي سبقتها، حيث يبدو واضحاً أن

نظم البريد الإلكتروني الموجودة حالياً لم تضطر إلى إجراء أي تغييرات في بنيتها لتتداول رسائل مايم، كما أن عمليات مراجعة إصدارات مايم، وعمليات تسجيل أنواع رئيسية وفرعية جديدة، وعمليات تبديل النظام أو تغييره كانت قد صممت جميعها لتتوافق بشكل طبيعي داخل الإجراءات القياسية للإنترنت، بل الأهم من ذلك كله أن مايم لا تحاول حل جميع قضايا أنواع البيانات، وليس من الصعوبة جعلها متكاملة مع أي تطبيقات محددة، يضاف إلى ذلك أنها تعمل على توفير مجموعة مرنة من الخدمات التي يمكن استخدامها في سياقات أو ظروف كثيرة متفاوتة، وأخيراً يمكن القول إن انتصارها الكبير قد احتل مكانة في عالم الويب، ذلكم التطبيق الذي لم يكن له وجود عندما ظهرت المايم.

الأنواع الهيكلية (البنائية) Structural Types :

عادة ما يبدو أي عنصر item للمستفيد وكأنه شيء واحد، دون إبداء أي اهتمام بالتمثيل الداخلي لهذا الشيء. كما ستشير أي ببليوجرافية أو كشاف وبشكل طبيعي لمثل هذه العنوان على أن الأمر ليس بهذه البساطة، حيث يمكن أن يكون التمثيل الداخلي بالشكل الذي اختزن به العمل في المكتبة الرقمية معقداً. وعليه فإن ما وراء البيانات البنائية تستخدم لتمثيل مكونات هذا العمل أو عرضها، وتحديد العلاقات القائمة بين هذه المكونات، مع ملاحظة أن اختيار ما وراء البيانات البنائية الخاصة بفئة محددة من المواد يستحدث نموذجا كائنياً.

ومما تجدر الإشارة إليه أن الفئات المختلفة من الكائنات تتطلب بالضرورة نماذج كائنية مختلفة، مثل النص بلغة الترميز العامة SGML،

وكائنات الويب، وبرامج الحاسب، والتسجيلات الصوتية المرقمنة. على أن تتولى القواعد والاتفاقيات الموجودة ضمن كل فئة من هذه الفئات وصف الكيفية التي يتم بها تنظيم المعلومات كمجموعات من الكائنات الرقمية؛ فهناك - على سبيل المثال - القواعد الخاصة بوصف كيفية تمثيل التسجيلات الصوتية المرقمنة أو عرضها، وبالنسبة لكل فئة على حدة، فتكون هناك قواعد محددة تصف كيفية تمثيل المواد في المكتبة، وكيفية تجميع المكونات كمجموعة من الكائنات الرقمية، كما تصف البنية الداخلية لكل عنصر، وما وراء البيانات المرتبطة به، كما تصف اتفاقيات تسمية الكائنات الرقمية. ولتمييز هذه الفئات يستخدم ما يعرف بالنوع البنائي أو الوسيمة البنائية structural type.

لقد احتلت نماذج الكائنات للبرامج الحاسوبية جزءاً قياسياً من عالم التحسيب على مدار سنوات عديدة. فالبرنامج الآلي الضخم يتكون عادة من العديد من ملفات البرامج والبيانات التي تضمنها بنية مركبة وتربطها علاقات متداخلة، وهذه العلاقات توصف في تركيبة بيانات مستقلة تستخدم في تجميع البرنامج وبنائه، وهذا ما يعرف في برنامج اليونكس بملف العمل make file.

وفي هذا الصدد يجب التمييز بين الأنواع البنائية، والأنواع أو الأجناس الأدبية genre، فمن الملائم بالنسبة لعملية استرجاع المعلومات أن يتم تقديم ما وراء بيانات وصفية تضطلع بمهمة وصف النوع أو الجنس الأدبي فهناك - على سبيل المثال - بعض أنواع للموسيقى الشعبية منها موسيقى الجاز jazz، وموسيقى البلوز blues، والراب Rap، والروك Rock، وإذا كان تحديد

الجنس الأدبي يعد وسيلة طبيعية ومفيدة كذلك لوصف المواد بالنسبة لعمليات البحث، ولغيرها من الأغراض الببليوجرافية الأخرى، فإن التجميعات أو التصنيفات الأخرى تعد ضرورة بالنسبة لعملية إدارة المكتبات الرقمية اللامركزية أو الموزعة، فعلى الرغم من أن الأفلام السينمائية، والأفلام الوثائقية، وأفلام الفيديو التعليمية، تعد من حيث نوعها مختلفة تمام الاختلاف بعضها عن بعض، فإن مناظراتها المرقمنة قد يتم تشفيرها بالطريقة نفسها، ومعالجتها بالإجراءات نفسها تقريباً، كما يمكن أن يكون لها النوع البنائي نفسه. وعلى العكس من ذلك، قد يكون هناك نصان من النوع أو الجنس الأدبي نفسه، ولكن سيكون لكل منهما نوع بنائي ونموذج كائني مختلفان تمام الاختلاف عندما يتم عرض أحدهما في صيغة اللغة المعيارية الموحدة، وعرض الآخر في صيغة بي دي إف.

وفي اللوحة رقم (١٢-٨) وصف لأحد النماذج الكائنية الخاصة بوصف الصور التي يتم أخذها عن طريق المسح الضوئي. وبالرغم من أن هذا النموذج كان قد تم تطويره لعرض الصور الفوتوجرافية المرقمنة، إلا أن النوع البنائي نفسه يمكن استخدامه مع أي صورة خرائطية bit-mapped، فعلى سبيل المثال، من المعروف أن الخريطة، والملصق الإعلاني، والإعلانات المسرحية، والرسوم البيانية الفنية، وبطاقات لعبة البيسبول تمثل محتويات مختلفة، لكنها جميعاً يتم اختزانها ومعالجتها في الحاسبات بالتركيبة نفسها. وترى الاتجاهات الفكرية المعاصرة أن المجموعات المعقدة بالمكتبات الرقمية يمكن تمثيلها في عدد قليل من الأنواع البنائية، وقد رؤيَ أن ما يقل عن عشرة أنواع بنائية ستكون كافية لوصف جميع فئات المواد

التـي يجـري تحـويلها فـي مكتـبة الكونجـرس، وكـان مـن بـين هـذه المـواد :
الـصور المـرقمنة، ومجموعات من صـور الـصفحات [العادية]، ومجموعات مـن صـور الـصفحات مـع الـنص المهـيأ بـصيغة اللغـة المعـيارية المـوحدة، والتسجيلات الصوتية المرقمنة، وتسجيلات الفيديو المرقمنة.

اللوحة رقم (١٢ـ ٨)

نموذج كائني لصور مأخوذة عن طريق المسح الضوئي

فـي عمـل مـشترك مـع مكتـبة الكونجـرس، قـام أعضـاء مـبادرات الـبحث الوطنــي بتطويـر نمـاذج كائنـية لـبعض أنـواع المـواد المـرقمنة عـن طـريق بـرنامج المكتـبة الـرقمية الوطنـية. وقـد كـان الـنموذج الأول خاصًّـا بالـصور المرقمنة، مثل الـصور الفوتوجـرافية المأخـوذة عـن طـريق المسح الـضوئي. وعند تحويل كل صورة من هذه الصور كان يتم عمل عدة نسخ رقمية منها، واحـدة عالـية الـنقاء لأغـراض الحفـظ الأرشـيفي، وأخـرى متوسـطة الـنقاء لأغراض الإتاحـة عبـر الـويب، وثالـثة منمنمة أو أكثـر من ذلـك، وكـان لكـل نـسخة من هذه النسخ مـا وراء البيانات الخاصـة بها، بالإضافة إلى ما وراء البيانات التي تشترك فيها جميع هذه النسخ.

وقـد تمـثل أول قـرار اتُّخِذ بـشأن تطويـر الـنموذج الكائنـي فـي اختزان ما وراء البيـانات الببليوجـرافية التـي تـصنف المحـتوى فـي تـسجيلات منفـصلة، وأن يـتم اختـزان كـل مـن مـا وراء البـيانات البيانـية (التـي تـربط بـين الـنسخ المخـتلفة) ومــا وراء البــيانات الإداريــة (التــي تــستخدم لأغــراض إدارة الوصـول) مـع الـصور الممـسوحة ضـوئياً، وأن تكـون جـزءاً مـن الـنموذج الكائنـي.

وبمـا أن النـسخ الـرقمية المـستقلة قـد تـستخدم بـشكل مـستقل فـي بعـض المناسبات، فـإن كل نـسخة يتم اختـزانها ككـائن رقمـي مستقل، كمـا أن هـذه النـسخ يتم تعـريفها أو تحديد هويـتها عن طـريق المعـالج (الهانـدل) handle، ويكون لهـا عنـصران همـا: الـصورة المـرقمنة، ومـا وراء البـيانات. وهـناك كـائن رقمـي آخر يعرف "بما وراء الكائن meta-object" يستخدم في إحضار النسخ بعضها مع بعض. وتتمثل الوظيفة الرئيسية لما وراء الكائن في قيامه بتقديم قائمـة من الإصدارات المختلفة، كمـا يتضمن هذا الكـائن كذلك حقلاً لنوع البيانات data-type field يـصف وظـيفة كل صـورة، بمعنى أنه يوضح أن هذه الـصورة مختـزنة لأغـراض الـرجوع إلـيها (الوظـيفة المـرجعية)، أو للإتاحـة عبر الإنتـرنت، أو كصورة منمنمة. كما يشتمل ما وراء الكائن الذي لــه معالجـة (handle) خاصــة بـه، علـى مـا وراء البـيانات التـي تـشترك فـيها جميع نسخ الصورة.

نماذج الكائنات لأغراض التشغيل المتداخل :

تـتطور نمـاذج الكائـنات ببطء، ومن الـنادر تحديد الكيفية التـي يجـب أن تُجمّـع بهـا الأجـزاء المتـنوعة بـناءً علـى عـدد قلـيل مـن القـواعد الـصارمة؛ فالقـرار يعتمد علـى الـسياق، وعلـى الكائـنات المحددة، وعلى نوع المحتوى، وأحيانـاً علـى المحتوى الفعلـي. فبعد ظـهور الطـباعة مثلاً، استغرقت عملية بـناء بنية الكتاب أو تشكيله عدة عقود لتـتطور وتصل على النحو الذي يبدو علـيه الآن والـذي نعـرفه بـه، بمـا فـيه مـن الـصفحات التمهـيدية والفـصول والأشـكال والكـشافات. ومــن غيــر المــدهش أن نجـد هــناك الآن قلــة مــن الاتفاقيات الخاصـة بنماذج الكائـنات وبمـا وراء البيانات البنائية في المكتبات

الرقمية ـ فما وراء البيانات البنائية لا تزال في المرحلة التي تقوم فيها كل مكتبة رقمية بتجريب مواصفاتها الخاصة بها، ومحاولة إجراء التعديلات التي تدعو إليها ضرورة التطوير.

فمن الواضح أن هذا الوضع يمثل عقبة لعملية التشغيل المتداخل، حيث إن البرنامج الخاص بالعميل لن يتعرف إلى ما وراء البيانات البنائية المستخدمة في تخزين الكائن الرقمي في أحد المستودعات المستقلة. ومع ذلك، فإن [برامج] الحاسبات العميلة لا تحتاج إلى التعرف إلى التفاصيل الداخلية الخاصة بكيفية اختزان المستودعات للكائنات، وإنما عليها أن تعرف الوظائف التي يمكن للمستودع أن يقدمها فقط.

وانظر مثلاً إلى عملية اختزان المواد المطبوعة التي تم تحويلها إلى الصيغ الرقمية، ستجد أن عنواناً واحداً في مجموعة إحدى المكتبات الرقمية يتكون من مجموعة من صور الصفحات، وإذا تم تحويل المادة كذلك إلى صفحات مرمزة عن طريق اللغة المعيارية الموحدة، فسيكون هناك عملية الترميز بهذه اللغة، ويكون هناك تعريف بوسيمة نوع الوثيقة data type document، ونماذج عرض الصفحات style sheets، وما يرتبط بكل ذلك من معلومات. وفي داخل المستودع، تقوم ما وراء البيانات البنائية بتحديد طبيعة العلاقة فيما بين المكونات، غير أن برنامج العميل لا يحتاج أن يعرف هذه التفاصيل. وقد تعرفنا في سياق الفصل الثامن على برنامج واجهة المستفيد، الذي يقوم بمعالجة سلسلة متتابعة من صفحات الصور، وقد تضمنت الوظائف على عمليات عرض صفحات محددة، والانتقال إلى إحدى الصفحات التي تحمل رقماً معيناً. ويكون بمقدور

واجهة المستفيد التي تعرف الوظائف الداعمة لمثل هذا النوع البنائي من المعلومات وإحضار المادة إلى المستفيد بدون معرفة كيفية اختزانها في المستودع.

البث أو النشر Dissemination :

نادراً ما يكون الشكل الذي اختزنت به المعلومات في المكتبة الرقمية هو نفسه الشكل الذي توصل به إلى المستفيد، وفي نموذج الويب للإتاحة، يتم استنساخ المعلومات من الخادم إلى حاسب المستفيد، حيث يتم هناك إحضارها للاستخدام من قبل المستفيد، وبالمثل، فإن عملية الإحضار تأخذ شكل عملية تحويل البيانات التي يتضمنها الملف إلى الصورة التي تظهر على الشاشة a screen image، باستخدام (الأبناط) والألوان المناسبة، وعن طريق إدخال الصورة أو تضمينها في وسائل عرض على النوافذ والقوائم والأيقونات.

ويمكن أن تكون عملية الحصول على المعلومات المخزنة وعرضها للمستفيد في المكتبة الرقمية أكثر وأشد تعقيداً مما يحدث في نموذج الويب. فأقل ما يحدث أن البنية الفنية العامة يجب أن تسمح لبرامج الحاسب بمعالجة المعلومات المختزنة على الخادم قبل إرسالها إلى حاسب العميل، أو معالجتها على [الحاسب] العميل قبل عرضها للمستفيد. وتجدر الإشارة إلى أن الأمر لا يتطلب أن يتم اختزان البيانات التي أحضرت بصراحة على الخادم، فهناك العديد من الحاسبات تقوم بتشغيل برامج آلية على مجموعة البيانات المختزنة، لتقوم بالتقاط معلومات معينة، وتقديمها للمستفيد. وهذه المعلومات قد تكون مجهزة تماماً من قبل الخادم، بيد أنها في الغالب ترسل أو تبث كملف تم تجهيزه [من قبل] بلغة ترميز النصوص الفائقة أو في بعض

الصيغ الوسيطة التي يمكن لحاسب المستفيد أن يتعرف إليها.

وتشتمل الأشكال الهامة لعملية النشر على عملية التفاعل المباشر بين العميل والكائنات الرقمية المختزنة والسيول المتدفقة والمستمرة من البيانات. وعندما يقوم المستفيد بالتفاعل مع المعلومات، يجب أن يدرك أن وصوله إلى هذه المعلومات قد لا يكون قاصراً عليه، وإنما يمكن أن يكون هناك من يستطيع الوصول إلى هذه المعلومات. وتعد ألعاب الفيديو خير مثال على ذلك، كما أن هذا الأمر ينطبق بحق على أي مادة تفاعلية، كما هو الحال عند محاكاة أحد المواقف الفعلية. فعملية الوصول إلى مثل هذه المعلومات تتكون من سلسلة من التفاعلات التي يتم توجيهها بواسطة تحكم المستفيد، أو بواسطة ما وراء البيانات البنائية الخاصة بالكائنات الفردية.

وغالباً ما يكون لدى برنامج العميل خيارات في البث، وقد يتحدد الاختيار بناءً على الإمكانيات التي يمتلكها المستفيد، فعلى سبيل المثال قد يجبر المستفيدون العاملون على الشبكات البطيئة، في بعض الأحيان، على إلغاء عرض الصور والاكتفاء بعرض النصوص على صفحات الويب. وقد يتحدد الاختيار بناء على إمكانيات البرمجية الموجودة على حاسب المستفيد، فعلى سبيل المثال، إذا كان النص متاحاً بصيغتين إحداهما بلغة ترميز النصوص الفائقة والأخرى بصيغة بوست سكربيت، فإن المستفيد الذي لا يتوافر لحاسبه البرمجية المتوافقة مع عرض نصوص البوست سكربيت يختار أسلوب البث المتوافق مع صيغة لغة ترميز النصوص الفائقة. بل إن تحديد أسلوب البث قد يعتمد على أسباب غير فنية، كرغبات المستفيد وراحته، فقد يفضل المستفيد رؤية نسخة قصيرة من برنامج الفيديو بدلاً من رؤية النسخة الكاملة.

إن أحد أهداف نماذج الكائنات أن تتيح للمستفيد أنواعاً متعددة من خيارات البث، كما تهدف كثير من مشروعات البحث إلى تمكين الحاسب العميل بشكل تلقائي من استكشاف عدد أساليب البث المتاحة، واختيار الأسلوب الأكثر فائدة في الوقت المناسب. ومن الناحية المثالية، عادة ما يتمثل أحد الخيارات في أسلوب البث الافتراضي - default dissemination - أي أن ما يستقبله المستفيد يكون هو الخيار ما لم يكن هناك خيار آخر. وهناك خيار آخر يتمثل في الملخص القصير short summery الذي يكون عادة عبارة عن سطر واحد من النص (وربما يكون مجرد اسم المؤلف أو عنوان هذا النص)، ولكن قد يكون هناك شيء أحسن أو أغرب من ذلك، كما هو الحال بالنسبة للصور المنمنمة، والمقطع القصير من صور الفيديو. ومهما تكن الطريقة المختارة، فإن هذا الملخص يتمثل في إتاحة كمية من المعلومات للمستفيد بشكل يساعده على التعرف إلى هوية الكائن [الرقمي] ووصفه. ومن نافلة القول إن هذا المجال لا يزال مجال بحث، وما هنالك سوى أمثلة قليلة وحديثة على مثل تلك النظم المستخدمة في الحياة العملية.

الفصل الثالث عشر
المستودعات والأرشيفات

يتناول هذا الفصل أساليب اختزان المواد الرقمية في المستودعات وحفظها لمدة طويلة، كما يتناول البروتوكولات التي توفر إمكانية الوصول إلى تلك المواد المختزنة. وعلى اعتبار أن التخزين لمدة طويلة هو أحد العناصر الرئيسية التي تقوم عليها المكتبات الرقمية، فقد يبدو غريباً للوهلة الأولى أننا أرجأنا تناول هذه الموضوعات الهامة إلى هذا الجزء الأخير من الكتاب، إلا أن لدينا مبررات لذلك؛ حيث ينصب تركيزنا في هذا الكتاب على ما هو موجود بالفعل في وقتنا الحاضر. ومع أننا عرضنا لموضوعات بحثية كلما وجدنا ذلك ضرورياً، إلا أن معظم مناقشاتنا كانت منصبة على النظم المستخدمة في المكتبات في وقتنا الحاضر. وفي حقيقة الأمر، ليس هناك اتفاق كامل في وجهات النظر حول القضايا التي سنعرضها في هذا الفصل حتى الآن، فباستثناء خادم الويب، هناك شبه إجماع حول مستودعات المكتبات الرقمية، كما أن الأرشفة الرقمية أو الحفظ الرقمي لا يزال مجالاً جديداً. وإن كانت قد ظهرت بوادر على بدء فهم الاحتياجات المرتبطة بهذه العملية، فإن أساليب العمل لم تتبلور بعد.

المستودعات :

المستودع هو نظام حاسوبي آلي تتمثل وظيفته الأولى في اختزان المواد الرقمية لاستخدامها في مكتبة ما. وعلى هذا الأساس فالمستودعات هي بمثابة أرفف الكتب في المكتبة الرقمية. وهذه المستودعات قد تكون ضخمة

وقد تكون صغيرة، تختزن ملايين من الكائنات الرقمية أو تقتصر على كائن رقمي واحد فقط. وفي بعض الحالات يمكن اعتبار أية وسيلة متنقلة تشتمل على قليل من الكائنات الرقمية مستودعاً، إلا أن معظم المستودعات تتألف من نظم حاسوبية آلية ضخمة تختزن المعلومات في نظام ملفات أو في قواعد بيانات وتتيحها إلى العالم عبر واجهة محددة ومعروفة.

خوادم الويب :

تُعد خوادم الويب النمط الأكثر شيوعاً للمستودعات في وقتنا الحاضر. وهناك شركات عديدة توفر هذه الخوادم الجيدة. وتكمن الفروق الرئيسية بين منتجات هذه الشركات في البرامج المصاحبة لهذه الخوادم مثل البريد الإلكتروني، وبرامج التكشيف، وأنظمة الأمن، وآليات التسديد المالي الإلكتروني، وغيرها من الخدمات الأخرى للشبكة.

أما فيما يتعلق بالمكتبات الرقمية، فإن خوادم الويب تؤدي مهام مقبولة بتكاليف منخفضة، وقد ساهمت هذه الخصائص في تقبل هذه الأجهزة على نطاق واسع وظهور مستوى يكاد يشكل قاعدة للتشغيل المتداخل. وتدين الويب بالفضل في جانب كبير من نجاحها إلى بساطتها وسهولتها، وما الحاسبات الخادمة إلا جزء من هذا النجاح، ومع ذلك يعد جانب من هذه البساطة التي تميز هذه الأجهزة غير ملائمة أو غير مريحة للقائمين على تنفيذ المكتبات الرقمية. فالحاسبات الخادمة لا تدعم سوى نموذج كائني واحد فقط: وهو نظام الملف الهرمي الذي تنتظم فيه المعلومات في ملفات منفصلة، ومعالجتها لهذه الملفات تعد معالجة غير كافية؛ حيث يتم تلقي كل

رسائلها ومعالجتها، وأخيراً نسيانها.

اللوحة رقم (١٣ـ١)

خوادم الويب

خادم الويب هو برنامج آلي تتمثل مهامه في اختزان الملفات والاستجابة للتساؤلات الواردة من بروتوكول نقل النص الفائق (HTTP) والبروتوكولات الأخرى المصاحبة له. ويتم تشغيل خادم الويب على حاسب آلي مرتبط بالإنترنت، ويمكن أن يكون هذا الحاسب مخصصاً لهذا الغرض فقط، أو أن يكون جهازاً مشتركاً يقوم بإجراء تطبيقات أخرى، أو أن يكون حاسباً شخصياً يتيح موقعاً صغيراً على الويب.

وفي قلب خادم الويب هناك عملية تسمى " httpd " [الحارس][١]، مع ملاحظة أن الحرف d الأخير يشير لكلمة daemon التي تعني شيطان. والحارس هو برنامج يعمل بشكل دائم، إلا أنه يبقى ساكناً معظم الوقت حتى تصله رسالة لمعالجتها. ويعمل بروتوكول نقل النص الفائق على قمة بروتوكول التحكم بالإرسال عبر الإنترنت TCP. ويوفر بروتوكول التحكم بالإرسال منافذ (عناوين) عديدة لكل حاسب، بحيث يرتبط الخادم بأحد هذه المنافذ، وهو المنفذ (رقم ٨٠) في العادة، وإن كان يمكن بالطبع تخصيص منافذ أخرى. وعندما تصل رسالة إلى هذا المنفذ، يقوم بتمريرها إلى الحارس، ومن ثم يبدأ الحارس في التعامل مع هذه الرسالة لتوزيعها، ويترقب في الوقت نفسه وصول رسائل أخرى. وبهذه الطريقة يمكن تداول

(١) اقتراح من المترجمين.

رسائل عديدة في الوقت نفسه من دون إشغال الحارس بتفاصيل معالجتها.

ويتحكم بروتوكول نقل النص الفائق بشكل جيد في المعالجة الفعلية التي يجريها خادم الويب للويب. ولم تكن أوائل خوادم الويب تقوم بأكثر من تنفيذ أمر "توصّل إلى أو احصل على get"، وهذا الأمر يعني تلقي رسالة تحتوي على محدد موحد لموقع المصدر URL من أحد الحاسبات العميلة، ويقوم هذا المحدد الموحد بتحديد ملف مختزن في الخادم، ومن ثم يسترجع الخادم الملف ويعيده إلى العميل. وبهذا تنتهي علاقة بروتوكول نقل النص الفائق بتلك الرسالة.

ومع ازدياد إمكانات بروتوكول نقل النص الفائق، وازدياد أعداد مواقع الويب أصبحت خوادم الويب أكثر تعقيداً. فعلاوة على تنفيذ مجموعة كاملة من أوامر بروتوكول نقل النص الفائق؛ أصبحت الأنظمة الخادمة تدعم نصوص واجهة بوابة المرور الموحدة CGI، والامتدادات الأخرى. ومن المتطلبات التي ما زال يتعين على الخوادم (ومتصفحات الويب كذلك) تلبيتها مواصلة دعم الإصدارات القديمة من بروتوكول نقل النص الفائق، بحيث يجب أن تكون هذه الخوادم جاهزة لتلقي أي رسائل بأي إصدار من هذا البروتوكول ومعالجتها على نحو سليم. وعلاوة على ذلك، تواصل الويب إضافة إمكانات الأمن أو الحماية. كما أن الإصدار (رقم ١,١) من هذا البروتوكول يحتوي كذلك على وصلات دائمة تسمح بمعالجة أوامر بروتوكول نقل النص الفائق عبر وصلة واحدة لبروتوكول التحكم بالإرسال.

وتحتاج المواقع الكبيرة على الويب إلى أكثر من حاسب نظراً لأنشطتها

الكثيرة. وهناك وسائل عديدة تستخدم للمشاركة في تحمل العبء، منها طريقة مباشرة تتمثل في إعادة نسخ البيانات على عدة أنظمة خادمة متطابقة، وتعد هذه الطريقة ملائمة ومريحة عندما يكون عدد التساؤلات كبيراً، وحجم البيانات متوسطاً. وهناك أسلوب آخر يسمى "خدمة اسم النطاق الدوارة DNS round robin" يستخدم لموازنة الحمل من خلال توسعة نظام اسم النطاق بحيث يسمح له أن يشير إلى مجموعة من الحاسبات ذات عناوين مختلفة لبروتوكول الإنترنت، وعلى سبيل المثال، يشير اسم النطاق "www.cnn.com" إلى مجموعة حاسبات، يوجد بكل واحد منها نسخة من موقع الويب الخاص بشبكة سي إن إن CNN، وعندما يدخل المستخدم إلى هذا الموقع يقوم نظام اسم النطاق باختيار أحد الحاسبات لخدمة الاستفسار، إلا أن استنساخ الموقع لا يكون مفيداً عندما يكون حجم البيانات ضخماً، أو إذا كانت البيانات سريعة التغير. وتعد خدمات بحث الويب مثالاً جيداً على ذلك؛ فبعض نظم البحث في الويب تستخدم حاسبات منفصلة لتنفيذ عملية البحث أو إجرائها، ثم تجميع الصفحات التي سترجع أو تقدم إلى المستخدم، هذا فضلاً عن استخدامها لأغراض عرض الإعلانات.

المستودعات المتطورة :

على الرغم من أن خوادم الويب تُستخدم على نطاق واسع، إلا أن هناك أنواعاً أخرى من أنظمة التخزين تُستخدم كمستودعات في المكتبات الرقمية. ففي معالجة بيانات قطاع الأعمال، تعد قواعد البيانات الارتباطية relational databases هي الطريقة المقننة في إدارة الكميات الكبيرة من البيانات. وتبنى قواعد البيانات الارتباطية على أساس نموذج كائني يتألف من قوائم البيانات

من جهة، وطبيعة العلاقات الموجودة بينها من جهة أخرى. وهذه العلاقات تسمح بجمع بيانات من قوائم مختلفة بطرق عديدة؛ فالقوائم وحقول البيانات الموجودة داخل قاعدة البيانات الارتباطية يتم تعريفها من خلال خطة وقاموس للبيانات. وتعد قواعد البيانات الارتباطية جيدة في إدارة كميات كبيرة من البيانات عندما تتم في سياق بنية محددة أو معرّفة بشكل جيد. وعليه فإن العديد من دور النشر الكبرى تعتمد على قواعد البيانات الارتباطية في إتاحة موادها، مدعومة بخوادم ويب تقوم بتوفير واجهات تربط بين المجموعات والمستفيدين.

أما بالنسبة لفهارس المكتبات الرقمية وكشافاتها فإنها عادة ما تتاح على نظم البحث التجارية، بحيث يكون لهذه النظم مجموعة من الكشافات التي تحيل إلى الكائنات الرقمية. وفعلياً يكون لهذه النظم نموذج مرن ومتقدم لتكشيف المعلومات، ونموذج آخر بدائي للمحتوى الفعلي. وكانت بدايات العديد من هذه النظم على شكل نظم النصوص الكاملة، وكان أقوى ما تقوم به هو توفير إمكانية لاسترجاع أحجام كبيرة من النصوص. وقد أضافت بعض النظم الأخرى إمكانات التغذية الراجعة ذات الصلة بالموضوع، بالإضافة إلى عمليات البحث الحقلي، وإمكانات أخرى، متطلعة بذلك إلى زيادة قدراتها الوظيفية وبالتالي زيادة مبيعاتها.

إن قواعد البيانات الارتباطية ونظم البحث التجارية توفر أدوات جيدة لتحميل البيانات، وتوثيقها، ومعالجتها، وحمايتها. كما تتسم عمليات إدارة الوصول بالدقة، وتتوافر كذلك الخدمات ذات الأهمية في التطبيقات التجارية (مسيرات مراجعة الحسابات التجارية). وهناك اتجاه واسع النطاق في

صناعة نظم قواعد البيانات لإضافة إمكانية بحث النصوص الكاملة، وأن تعمل نظم البحث على توفير بعض أجزاء ومكونات نموذج قاعدة البيانات الارتباطية، وقد تعد هذه المميزات الإضافية مفيدة، إلا أنه حتى الآن لم تتمكن أي شركة من إنتاج نظام يجمع أفضل ما في هذين المنهجين.

وبالرغم من أن بعض المكتبات الرقمية تستخدم قواعد البيانات الارتباطية بنجاح، فإن النموذج الارتباطي للبيانات ليس مرناً بما فيه الكفاية بالنسبة للنماذج الكائنية القوية الناشئة. وتجمع المكتبات الرقمية الرائدة فيما بينها على أهمية الحاجة إلى مستودعات أكثر تطوراً. وفيما يلي مجموعة من المتطلبات الضرورية لمثل هذه المستودعات.

إمكانية إخفاء المعلومات Information hiding :

يجب إخفاء التنظيم الداخلي للمستودع عن الحاسبات العميلة، وأن تتوافر إمكانية إعادة تنظيم أي مجموعة من المجموعات أو تغيير وضعها الداخلي، أو نقلها إلى حاسب آخر من دون أن يكون لذلك أية تأثيرات على المجموعات الأخرى.

النماذج الكائنية Object Models :

يجب أن تدعم المستودعات مجموعة مرنة من النماذج الكائنية، مع تقليل القيود المفروضة على البيانات، وما وراء البيانات، والروابط الخارجية، والعلاقات الداخلية. ولا ينبغي أن يؤدي إدخال فئات جديدة من المعلومات إلى إحداث تغييرات جوهرية في بنية الجوانب الأخرى للمكتبة الرقمية.

البروتوكولات والصيغ المفتوحة Open protocols and formats :

يجب أن تتمكن الحاسبات العميلة من الاتصال بالمستودعات عن طريق

بروتوكولات البيانات وأنواعها، صيغ معرفة على نحو واضح، ويجب أن تسمح البنية الفنية للمستودع بإدخال تغييرات جوهرية على البروتوكولات كلما دعا تطويرها إلى ذلك. ولعل هذا أكثر ما ينطبق على عملية إدارة الوصول بصفة خاصة. وأخيراً يجب أن يسمح المستودع بتطبيق مجموعة من السياسات على جميع المستويات.

مدى الموثوقية وكفاءة الأداء Reliability and performance :

يجب أن يكون المستودع قادراً على تخزين كميات كبيرة للغاية من البيانات، وأن يكون موضعاً للثقة الكاملة، وأن يعمل جيدًا.

ما وراء البيانات في المستودعات :

تختزن المستودعات البيانات وما وراء البيانات. وما وراء البيانات هذه يمكن تصنيفها في ثلاث فئات عامة هي: ما وراء البيانات الوصفية، وما وراء البيانات البنائية أو الهيكلية، وما وراء البيانات الإدارية، كما قد تكون هناك حاجة إلى إيجاد محددات للتمييز بين عناصر الكائنات الرقمية فضلاً عن محددات وصف الكائنات نفسها. ويتطلب تخزين ما وراء البيانات في المستودعات إجراءات تتسم بالمرونة. وهناك عدة احتمالات لعملية اختزان ما وراء البيانات هذه، تتمثل فما يلي :

- ما وراء البيانات الوصفية، وتختزن غالباً في فهارس وكشافات تتم إدارتها خارج المستودع. ويمكن الاحتفاظ بهذه الفهارس والكشافات في مستودعات منفصلة، كما يمكن لهذه الفهارس والكشافات أن تمتد تغطيتها لتشمل مواد في عدة مكتبات رقمية مستقلة. أما المحددات فتستخدم في هذه الحالة لربط ما وراء البيانات مع البيانات ذات الصلة

بها.

- مـا وراء البـيانات البنائـية والإداريـة، وتختـزن غالـباً مـع الكائـنات الـرقمية، وغالـباً مـا تكون مـا وراء البيانات هـذه مضمنة فعلياً داخـل الكائن نفسه.
- بعض مـا وراء البيانات تشير إلى مجموعة من الكائنات، أما ما وراء البـيانات الإداريــة المـستخدمه لأغـراض إدارة الوصــول فـيمكن أن تـستخدم أيضاً مـع المـستودع بكاملـه أو مـع إحـدى المجمـوعات داخـل المستودع، كما تستخدم معينات البحث مع العديد من الكائنات.
- يمكـن اختـزان مـا وراء البـيانات ككائـنات رقمـية منفـصلة مـع إيجـاد روابط من الكائنات الـرقمية التـي تنطـبق عليها ما وراء البيانات هذه. وهـناك بعـض مـن مـا وراء البـيانات التــي لا تختـزن صـراحة، ولكـن يمكن إيجادها أو إنشاؤها عند الحاجة إليها.

ونظـراً لوجـود كـل هـذه الاحـتمالات، فـإن لكـل مكتـبة رقمـية أفكارهـا الخاصــة بـشأن اختـيار مــا وراء البـيانات ووصـفها، إلا أن تـبادل مــا وراء البـيانات يبقـى قـضية القـضايا بالنـسبة لعملـية التـشغيل المـتداخل. ومـا نظـام العمل أو إطـاره المسمى وارويك Warwick، الوارد في اللوحة رقم (١٣-٢) إلا شكل من أشكال التنظيم المقترحة للتعامل مع هذا الموقف المضطرب.

اللوحة رقم (١٣– ٢)

إطار عمل وارويك **The Warwick Framework**

يـرجع أصـل نظـام أو إطـار عمـل وارويـك إلـى بعـض الأفكـار التــي

تمخضت عنها ورشة العمل التي عقدت في رحاب جامعة وارويك في إنجلترا عام ١٩٩٦م. والهدف الأساسي لهذا النظام هو تنظيم ما وراء البيانات. فهناك مجموعة كبيرة من ما وراء البيانات - مثل ما وراء البيانات الوصفية كبيانات الفهرسة المقروءة آلياً MARC، و ما وراء البيانات الخاصة بإدارة الوصول، و ما وراء البيانات البنائية، والمحددات - يمكن أن تطبق على كائن رقمي واحد. وقد اقترح الأعضاء المشاركون في الورشة أن يتم تنظيم ما وراء البيانات في حزم مستقلة، كما هو الحال بالنسبة لحزمة بيانات دبلن كور Dublin Core، والحزمة الخاصة بالبيانات المكانية Geospatial. ولا شك أن لهذا الفصل (لما وراء البيانات) مميزات كثيرة في تيسير عملية التشغيل المتداخل، بمعنى أنه إذا كان كل من الحاسب العميل والمستودع قادرين على معالجة حزم من نوع معين، فإنهما سيكونان قادرين على العمل المتداخل أو المشترك إلى حد ما، حتى إن لم يكونا مشتركين في حزم ما وراء البيانات الأخرى التي يدعمها كل منهما.

ويتبنى كل من كارل لاجوز Carl lagoze من جامعة كورنيل، ورون دانيل Ron Daniel من مختبر لوس ألاموس الوطني، هذه الفكرة البسيطة، كما طورا طريقة رائعة لاستكشاف جميع عناصر الكائن الرقمي. وكانت أولى ملاحظاتهما أن ثمة غموضاً كثيراً في التمييز بين البيانات و ما وراء البيانات، وضربا مثلاً لذلك بتساؤلهما عن مدى اعتبار قائمة محتويات الكتاب جزءاً من محتواه، أم أنها تعد بمثابة ما وراء بيانات خاصة بمحتوى الكتاب؟ وأبانا أن مثل هذا التمييز ليست له أهمية في نظام وارويك. لأن كل

شيء يتم تقسيمه إلى حزم، ولا يتم التمييز بين البيانات وما وراء البيانات. أما الملاحظة الثانية فتمثلت في أنه ليس من الضروري اختزان كل حزمة على نحو صريح على أنها جزء من كائن رقمي؛ فما وراء البيانات الوصفية غالباً ما تختزن مستقلة وكأنها تسجيلة في أحد الفهارس أو في الكشافات، كما أن الشروط والأحكام التي تطبق على العديد من الكائنات الرقمية تختزن في أفضل حالاتها في تسجيلات ضبط أو تسجيلات تنظيمية منفصلة policy record، وليس بإدراجها ضمن كل كائن الرقمي. ويمكن أن يتم تنفيذ هذا الفصل عن طريق السماح بوجود حزم غير مباشرة، حيث تخزن الحزمة في أي مكان مناسب لها، مع وجود إشارة مرجعية لها في المستودع، وهذه الإشارة المرجعية قد تكون بمثابة مؤشر بسيط إلى موقع ما، أو قد يستعان ببرنامج آلي يؤدي تشغيله إلى جلب الحزمة عند الحاجة إليها.

إن الكائنات الرقمية من قبيل النوع البنائي الواحد عادة ما تتألف من مجموعة خاصة من الحزم. وهذا من شأنه إيجاد نموذج كائني يستخدم لأغراض التشغيل المتداخل بين العملاء والمستودعات لهذا النوع من الكائنات الرقمية. ومما تجدر الإشارة إليه أنه بالرغم من أن نظام وارويك لم يطبق على نحو صريح في أي نظام واسع النطاق، فإن الأفكار التي أثارها تبدو واعدة. فتقسيم المعلومات إلى حزم محددة بوضوح يبسط عملية وصف الكائنات الرقمية، كما يوفر المرونة اللازمة لعملية التشغيل المتداخل.

بروتوكولات التشغيل المتداخل :

تحتاج عملية التشغيل المتداخل إلى بروتوكولات يمكن للعملاء

استخدامها لإرسال رسائل إلى المستودعات، ويمكن للمستودعات بدورها استخدامها لإعادة إرسال المعلومات إلى العملاء. وعلى مستوى الأساس، هناك بعض العمليات اللازمة لتخزين المعلومات في مستودع وتوفير إمكانية الوصول إليها. ويتطلب استخدام نظم فعالة أن يكون العميل قادراً على اكتشاف بنية الموضوعات الرقمية. كما أن وجود أنواع مختلفة من الكائنات يتطلب توافر أساليب وصول مختلفة، وقد تتطلب عملية إدارة الوصول توافر ضمانات التثبت من الهوية أو إجراء مفاوضات بين العميل والمستودع. وعلاوة على ذلك، قد يرغب العملاء في بحث في الكشافات الموجودة داخل المستودع.

وفي الوقت الحاضر، يعد بروتوكول نقل النص الفائق (HTTP)، وهو بروتوكول الدخول إلى الإنترنت، أكثر البروتوكولات استخداماً في المكتبات الرقمية. وهناك بروتوكول آخر يستخدم على نطاق واسع هو بروتوكول تبادل البيانات الببليوجرافية،والمعروف بمعيار Z39.50.

اللوحة رقم (١٣- ٣)

بروتوكول نقل النص الفائق **(HTTP)**

تعد رسالة "توصل إلى أو احصل على get" في بروتوكول نقل النص الفائق أمرًا من أحد الحاسبات العميلة إلى أحد الحاسبات الخادمة لاسترجاع أية معلومات قام بتعريفها المحدد الموحد لموقع المصدر URL الذي تضمنته الرسالة. وإذا أشار المحدد الموحد لموقع المصدر إلى عملية ما تقوم بإنتاج بيانات، فإن البيانات التي ستنتجها هذه العملية ستكون هي البيانات

المسترجعة.

وتبدأ الاستجابة لأمر "توصل إلى" بترميز الحالة Status code، المكون من ثلاث تمثيلات رقمية، وقد تكون هذه الرموز مألوفة لمستخدمي الويب وكأنها رموز أخطاء، مثال ذلك الرموز ٤٠٤، الذي يسترجع في حالة عدم العثور على المصدر الذي تم تحديد عنوانه عن طريق المحدد الموحد لموقع المصدر. وفي حالة عدم وجود خطأ، تتبع رموز الحالة بمعلومات فنية (تستخدم أساساً لدعم الملقمات والذاكرات الفورية أو مساعدتها caches)، ثم بما وراء بيانات عن جسم الاستجابة response body. وتقوم ما وراء البيانات بإعطاء العميل معلومات عن عدد الوحدات الطولية الخاصة بنوع البيانات، ولغتها، وتشفيرها، والقيمة الاختبارية القصيرة المعروفة بالهاش hash، والتاريخ. ويستخدم العميل ما وراء البيانات هذه لمعالجة جسم الاستجابة، وهو الجزء الأخير من الرسالة الذي هو عادة الملف الذي يشير إليه المحدد الموحد.

وهناك نوعان آخران من رسائل بروتوكول نقل النص الفائق لهما ارتباط وثيق برسالة "توصل إلى"، هما: الرسالة الأمامية head message، وهي تطلب البيانات نفسها التي تطلبها رسالة "توصل إلى"، إلا أن جسم الرسالة نفسه لا يرسل، وهذه الرسالة تُفيد في اختبار روابط النص الفائق من حيث المصداقية، أو إمكانية الوصول، أو التعديلات الجديدة من دون الحاجة إلى نقل ملفات كبيرة. أما الرسالة الأخرى فهي رسالة المركز post message، وهي تستخدم لتوسعة حجم المعلومات التي يرسلها العميل إلى الخادم، وتستخدم رسائل المركز بشكل شائع من قبل العميل لإرسال إحدى

كتل البيانات، كما هو الحال في صيغ لغة ترميز النصوص الفائقة HTML، التـي يمكن معالجتها بعد ذلك عن طريق نص واجهة بوابة العبور الموحدة أو من خلال بعض التطبيقات الأخرى على الخادم.

ويتمثل الاستخدام الأساسي لبروتوكول نقل النص الفائق في استرجاع المعلومات من الخادم، ويمكن استخدامه لتغيير المعلومات على الخادم، كما تستخدم رسالة أخرى تعرف برسالة "ضع put" لاختزان معلومات معينة في أحد المحددات الموحدة؛ أما رسالة "ألغ delete" فتستخدم لإلغاء المعلومات. تجدر الإشارة إلى أن رسالتي "ضع"، و"ألغ" نادراً ما تستخدمان لأن الطريقة المعتادة لإضافة معلومات إلى خادم الويب تكون من خلال برامج منفصلة تعالج المعلومات الموجودة على الخادم، وليس من خلال رسائل بروتوكول نقل النص الفائق المرسلة من الخارج.

لقد أُدْخِلت تغييرات كثيرة على بروتوكول نقل النص الفائق منذ ابتكاره من أجل إتاحة الفرصة لإيجاد إصدارات مختلفة من هذا البروتوكول في الوقت نفسه، ولتعزيز أدائه على الإنترنت. كما يقوم بروتوكول نقل النص الفائق بالتعرف إلى أن هناك رسائل عديدة تُعالج من قبل كل من الملقمات أو الذاكرات الفورية caches. وتحتوي الإصدارات المتأخرة من هذا البروتوكول على مجموعة متنوعة من البيانات والخدمات التي تساند هذه التداخلات في معالجة الرسائل. وهناك أيضاً أنواع رسائل خاصة مثل رسالة "الخيارات options"، التي تسمح للعميل بطلب معلومات عن خيارات الاتصالات المتاحة، ورسالة "تتبع trace" التي تستخدم لتشخيص المشكلات وفحصها.

وبالرغم من الاتساع الذي شهده بروتوكول نقل النص الفائق مع مرور الأيام، فإنه لا يزال بروتوكولاً بسيطاً. وقد نجح مصمموه في مقاومة الضغوط الهادفة إلى إضافة مزيد من الخصائص إليه، كما تمكنوا في الوقت نفسه من إدخال تعديلات عليه لتحسين أدائه. وقد لا يتفق شخصان في الرأي حول ما يجب أن يقدم هذا البروتوكول، إلا أن هذا البروتوكول يبقى بلا شك أحد أكبر النجاحات في عالم الإنترنت.

البرمجة الموجهة للكائنات Object-oriented programming، والكائنات الموزعة :

يهدف أحد اتجاهات البحث إلى تطوير أبسط بروتوكولات المستودعات الممكنة لدعم الوظائف الضرورية. وإذا كان بروتوكول المستودع بسيطاً، فإن المعلومات عن أنواع الكائنات المعقدة يجب أن تكون مضمنة في الكائنات الرقمية، وهذا الأسلوب يعرف اختصاراً بسودا SODA، أي "الكائن الذكي والأرشيفات الصماء smart object, dumb archive".

وهناك العديد من المشروعات المتقدمة تقوم على تطوير بنيات فنية تستخدم المفهوم الآلي للكائنات الموزعة. وفي هذا السياق نجد أن كلمة "كائن object" لها معنى فني دقيق يختلف عما تعنيه الكلمة في سياق مصطلحي "الكائن الرقمي" و "الكائن المكتبي" المستخدمان في أماكن أخرى من هذا الكتاب. ففي مجال التحسيب الحديث، يعد الكائن جزءًا مستقلاً من شيفرة الحاسب والبيانات المصاحبة لها، والتي يمكن استخدامها وإعادة استخدامها في سياقات عديدة. وعادة ما تكون المعلومات والآليات بداخل الكائن مغلّفة encapsulated، مما يجعل جميع التفاصيل الداخلية تبدو مختفية.

وكل ما يعرفه العالم الخارجي عن فئة من الكائنات هو مجرد واجهة عامة تتكون من أساليب (أي العمليات التي تجري على الكائن) وبيانات فورية. وقد يتفاوت تأثير أسلوب معين من فئة إلى أخرى. فعلى سبيل المثال، في المكتبة الرقمية قد يكون لطريقة الإحضار أو الترجمة render معاني أو تفسيرات مختلفة لفئات مختلفة من الكائنات.

وبعد عقود من التطور، أصبحت لغات البرمجة الموجهة للكائنات، مثل لغتي : سي بلص بلص ++C، وجافا، مقبولة باعتبارها أفضل الوسائل المنتجة لبناء النظم الآلية. ولقد كانت القوة الدافعة وراء البرمجة الموجهة للكائنات هي تعقد عمليات التحسيب. يضاف إلى ذلك أن البرمجة الموجهة للكائنات تسمح بتطوير المكونات واختبارها على نحو منفصل، بحيث لا يتطلب تنقيحها لإصدارات متتابعة من النظام. وتعد شركة مايكروسوفت أحد المستخدمين الهامين للبرمجة الموجهة للكائنات. وهناك إصدارات عديدة من بيئة البرمجة الموجهة للكائنات أطلقتها شركة مايكروسوفت تعرف بـ OLE، و COM، و DCOM، و Active X، وكلها أشكال متنوعة للمفاهيم الأساسية نفسها.

إن الكائنات الموزعة تدعو لتعميم فكرة الكائنات على البيئة الشبكية. والفكرة الأساسية تتمثل هنا في أن الكائن الذي يجري تنفيذه على أحد الحاسبات يجب أن يكون قادراً على التفاعل مع كائن آخر بحيث يُنفذ على حاسب آخر عبر واجهته المنشورة، والمعروفة من حيث أساليب التعامل والبيانات الفورية التي تقدمها. وقد طورت شركات البرمجيات الرائدة (باستثناء شركة مايكروسوفت) معياراً للكائنات الموزعة يعرف باسم كوربا

CORBA [1]. ويوفر هذا المعيار لمطوري أنظمة نظم الحاسبات الموزعة وسائل الراحة البرمجية نفسها التي تقدمها البرمجة الموجهة للكائنات داخل حاسب آلي واحد.

وتتمثل الفكرة الأساسية لهذا المعيار في وجود وسيط طلب الكائن (Object Request Broker (ORB)). وعند إضافة هذا الوسيط إلى أحد برامج التطبيقات فإنه يقيم علاقات من نوع علاقة العميل / الخادم بين الكائنات. بحيث يتمكن العميل باستخدام الوسيط من إيجاد طريقة على خادم الكائن، وهذه الوسيلة قد تكون موجودة في الجهاز نفسه أو عبر شبكة. ويقوم الوسيط بتلقي الاتصال وإيجاد كائن يمكنه تلبية الطلب، ومن ثم يمرر المؤشرات إلى ذلك الكائن، ويوجد الوسيلة، ثم يسترجع النتائج. ولا يقتضي الأمر أن يكون العميل على علم بمكان وجود الكائن، أو لغته البرمجية، أو نظام تشغيله، أو أي جوانب أخرى من عناصر النظام التي لا تعد جزءاً من واجهة الكائن. وهكذا فإن الوسيط يعمل على توفير التشغيل المتداخل بين تطبيقات موجودة في أجهزة مختلفة تعمل في بيئات موزعة ومختلفة أيضاً.

إخفاء البيانات data hiding :

على الرغم من أن مفهوم الكائنات في سياق مجال التحسيب، يختلف عن مفهوم الكائنات الرقمية في سياق المكتبات، فإن هناك بعض السمات المشتركة بينهما، أما مصطلح "إخفاء البيانات" فهو مستمد من مجال البرمجة الموجهة للكائنات، لكنه ينطبق كذلك تماماً على الكائنات الرقمية في المكتبات. فعندما يدخل أحد العملاء إلى معلومات في أحد المستودعات فإنه

(١) اختصار لـِ Computing Object Request Broker Another (المترجمان).

يحتاج إلى معرفة طبيعة الواجهة التي يقدمها المستودع للعالم الخارجي، ولكنه لا يحتاج إلى معرفة كيفية تخزين المعلومات في المستودع. وفي سياق خادم الويب نجد أن الواجهة تتمثل في بروتوكول (بروتوكول نقل النص الفائق HTTP)، وخطة عناوين (المحدد الموحد لموقع المصدر URL)، ومجموعة من الصيغ وأنواع البيانات. أما في سياق مستودعات أخرى، فيمكن التعبير عن ذلك بمصطلحات البرمجة الموجهة للكائنات. وما يعتقد المستخدم أنه كائن رقمي واحد قد يكون مخزناً في المستودع على شكل مجموعة معقدة من الملفات، أو كمجموعة تسجيلات في قوائم، أو كائنات نشطة تنفذ عند الطلب.

إن إخفاء البنية أو التركيب الداخلي، وتوفير جميع إمكانات الوصول من خلال واجهة محددة على نحو جيد، ييسر عملية التشغيل المتداخل. ولا شك أن العملاء قد يستفيدون من عدم احتياجهم إلى معرفة التنظيم الداخلي للمستودعات. فقد يلجأ مستودعان إلى تنظيم معلومات متماثلة بأساليب مختلفة؛ حيث يمكن لأحدهما أن يختزن مسار الصوت وصور الفيلم الرقمي على أنهما كائنان رقميان منفصلان، في حين يختار الآخر اختزانهما على أنهما كائن واحد. ولذا يجب أن يكون برنامج العميل قادراً على إرسال طلب لبدء القراءة أو التشغيل غير مكترث بهذه الاختلافات الداخلية. كذلك تستفيد المستودعات من ذلك لأن التنظيم الداخلي يبقى شأناً داخلياً محضاً. وما يبدو للمستخدم على أنه كائن رقمي واحد، قد يكون في حقيقة الأمر صفحة في صيغة لغة ترميز النص الفائق مع صور مرتبطة بها مع تطبيقات

بلغة جافا. ومع طريقة إخفاء المعلومات، يمكن نقل الصور إلى موقع آخر أو التحويل إلى نسخة أو إصدارة جديدة بلغة جافا من دون أن يدري العملاء عن ذلك.

وتوفر كل من خدمة جستور JSTOR، وبرنامج المكتبة الرقمية الوطنية في مكتبة الكونجرس، صوراً منمنمة كإصدارات مصغرة scaled-down versions لصور أكبر. وقد قررت مكتبة الكونجرس إنتاج الصور المنمنمة مقدماً وتخزينها على أنها بيانات مستقلة. أما خدمة جستور فإنها لا تختزن الإصدارات المصغرة، وإنما تقوم بدلاً من ذلك باستخراجها عند الطلب من الأشكال المختزنة للصور الأكبر، وكلا الأسلوبين مقبولان. وهذه القرارات الداخلية يجب أن ألا تظهر على واجهة المستفيد، كما يمكن أن يتم تغييرها في أوقات لاحقة؛ فالنظم الداخلية تحتاج إلى معرفة أن المستودع قادر على تزويدها بصورة منمنمة لا إلى معرفة كيفية استخدامها.

النظم الموروثة Legacy system :

اتبعت المحاولات النمطية للتشغيل المتداخل مساراً يعتمد على معايير متفق عليها. وهذه السياسة تتمثل في إقناع جميع الهيئات بالاتفاق على مجموعة من المعايير الفنية والتنظيمية. أما بالنسبة للنظم المعقدة كالمكتبات الرقمية، فالمهمة شاقة، وهناك حاجة للمعايير لضبط العمل في مجالات كثيرة كبناء الشبكات، وأنواع البيانات، وأساليب التعريف، وسبل الأمن، وأساليب البحث والاسترجاع، والكشف عن الأخطاء، وتبادل دفع رسوم التكاليف، وكل معيار من المعايير الموجهة لهذه العمليات يتكون من أجزاء عديدة، تتناول

الإجراءات المرتبطة بعمليات الصياغة اللفظية والدلالية، وإجراءات تفحص الخطأ، وإعداد التوسعات، وغير ذلك. وإذا تم الاتفاق على مجموعة كاملة من المعايير، وإذا طبقت الهيئات هذه المعايير تطبيقاً كاملاً، فإننا قد نتمكن من تحقيق مستوى رائع للتشغيل المتداخل.

وإذا تأملنا واقع التطبيق العملي، نجد أن السرعة التي تسير بها خطى عملية المعايرة أبطأ من سرعة التطور التقني، إذا لم تتمكن هيئة واحدة من تطبيق المعايير على نحو كامل في أنظمتها، بل حتى قبل أن نبدأ هذا التطبيق الكامل تسارع الهيئات إلى تغيير أنظمتها للاستفادة من فرص جديدة أو لتفادي مشكلات جديدة. والسؤال الذي يطرح نفسه هنا فيما يتعلق بالتشغيل المتداخل هو: كيف تتمكن نظم من أجيال مختلفة أن تعمل بعضها مع بعض ؟، وإن كان يطلق على النظم القديمة في بعض الأحيان اسم "النظم الموروثة" - استخفافاً بها، فإن هناك كثيراً من النظم القديمة لا تزال تقوم بأعمال جيدة، ولذا يتعين دائماً أن تستوعب الخطط المستقبلية النظم والالتزامات الحالية.

الحفظ التاريخي أو الأرشيفي Archiving :

الأرشيفات هي المادة الخام المادة للتاريخ، وتتجسد مهمة إدارة الأرشيف الوطني والوثائق الجارية في الولايات المتحدة الأمريكية في حفظ الوثائق "طوال حياة الدولة". وعلى أقل تقدير يجب أن تكون الأرشيفات مجهزة للقيام بعمليات حفظ المعلومات المكتبية لفترات أطول من أي نظام آلي آخر موجود اليوم، وأطول أيضاً من أية وسيلة إلكترونية أو ممغنطة تمت

تجربتها. ولا شك أن عملية الحفظ الأرشيفي الرقمي مهمة ليست باليسيرة. وقد يكون طرح القضايا أسهل من علاجها، وهناك عدة أمثلة جيدة على ذلك. ويعد التقرير الصادر عن لجنة العمل الخاصة بالحفظ الأرشيفي للمعلومات الرقمية Task force on Archiving of digital information، الوارد في اللوحة رقم (١٣- ٥)، هو الأساس الذي يقوم عليه العمل الحديث في مجال الحفظ الأرشيفي الرقمي.

وتميز عملية الحفظ الأرشيفي التقليدي بين مصطلحي الحماية Conservationالذي يعنى بالحفاظ المادي على المواد، والحفظ أو الصيانة Preservation، الذي يسعى للمحافظة على المحتوى حتى في حالة تلف المادة الأصلية أو دمارها. وهناك أسلوبان مناظران لهما في سياق الحفظ أو الأرشفة الرقمية، هما: التحديث أو التنشيط refreshing الذي يهدف إلى المحافظة على التسلسل الدقيق للبتات، والنقل migration الذي يسعى إلى المحافظة على المحتوى بمستواه الدلالي أو اللفظي، ولكن من دون المحافظة على التسلسل الدقيق للبتات. وأول من أقر هذا التمايز هي لجنة العمل الخاصة بالحفظ الأرشيفي للمعلومات الرقمية، التي أوصت باتخاذ "النقل" أسلوباً أساسياً في عملية الأرشفة الرقمية.

ويتطلب التنشيط والنقل جهوداً متواصلة؛ فوثائق الأعمال تحفظ لفترات زمنية طويلة لأن وراءها فريقًا من الأفراد يأخذون أجوراً للمحافظة عليها وصيانتها، وهذا هو عملهم. وهؤلاء الأفراد يعتنون بالقضايا المرتبطة بهذه العملية مثل الأمن، والنسخ الاحتياطي، وتوفير البيانات على المدى الطويل. وقد تنبه الناشرون أيضاً إلى أن معلوماتهم الرقمية تعد إرثاً يمكن أن يدر

عليهم أرباحاً لعقود، إلا أنه، بالنسبة للعديد من المجموعات الرقمية، لا يوجد شخص مسؤول عن المحافظة على المعلومات لفترات تتجاوز مصلحته الآنية. وهناك بيانات قد تبدو اليوم غير ذات فائدة، ولكنها يمكن أن تقدر تقديراً كبيراً في المستقبل البعيد. ولكن الحفظ الأرشيفي لمثل هذه المعلومات يحتل مرتبة دنيا ـ مع الأسف ـ في قائمة أولويات كل فرد، وهي أول شيء يستغنى عنه عند تخفيض الميزانية.

اللوحة رقم (١٣ـ٤)

مشروع إنفوباص في جامعة ستانفورد **The Stanford InfoBus**

هذا المشروع الذي يعرف بإنفوباص Infobus هو أحد المشروعات التي مولتها مبادرة المكتبات الرقمية وكان يديره هكتور جارسيا ـ مولينا Hector Garcia -Molina، ثم تيري وينوجراد Terry Winograd، ثم أندرياس بيبك Andrwas Paepake. ويعالج هذا المشروع مشكلة التشغيل المتداخل بين النظم العاملة، من خلال تقبل الأنظمة الموجودة كما هي، دون محاولة إرساء معايير جديدة أو تغيير النظم القديمة. والفكرة الأساسية لهذا المشروع تتمثل في إنشاء ملقمات خدمة مكتبية تعرف بكائنات كوربا CORBA لتمثل الخدمات الإلكترونية المباشرة. وهذه الملقمات تقوم بالاتصال بالخدمات الموجودة من خلال أية وسيلة اتصال من الوسائل التي تدعمها، ثم تقوم بتحويل الرسائل إلى واجهات قياسية معرّفة لكوربا، فعلى سبيل المثال، حين يرغب أحد العملاء لديه واجهة بحث يدّعمها معيار Z39.50 في البحث في إحدى خدمات البحث المباشرة مثل دايالوج

DIALOG، فهذا الأمر يتطلب استخدام ملقمين، أحدهما للترجمة بين بروتوكول البحث Z39.50، ونموذج الإنفوباص InfoBus، أما الآخر فللترجمة بين واجهة دايالوج ونموذج إنفوباص. وباستخدام هذين الملقمين يستطيع العميل البحث في دايالوج DIALOG بالرغم من وجود واجهتين مختلفتين.

وربما تكون أكثر أدوات الإنفوباص أهمية هي تلك التي تدعم معيار Z39.50، وقد طورت جامعة ستانفورد ملقماً يمكّن الحاسبات التي يدعمها معيار Z39.50 من الاتصال بخدمات البحث التي لا يدعمها هذا المعيار؛ بحيث يمكن للمستفيدين تقديم تساؤلات البحث إلى هذا الملقم عن طريق أي واجهة ربط مصممة للاتصال بخادم يدعمه معيار Z39.50. من ثم يقوم الملقم من خلال الإنفوباص بتمرير طلب البحث إلى أي مصدر من المصادر المتاحة الخاصة بالإنفوباص، حتى وإن كانت هذه المصادر لا تدعم معيار Z39.50، وبعد ذلك يقوم الملقم بتحويل النتائج إلى صيغة يمكن فهمها من قبل الحاسبات العميلة التي يدعمها معيار Z39.50. وقام الباحثون في جامعة ميتشجان بعمل مواز لذلك؛ حيث قاموا بعمل ملقم آخر يساعد على إمكانية الوصول إلى جميع الحاسبات الخادمة التي يدعمها المعيار Z39.50 عن طريق الإنفوباص. وعلاوة على ذلك تبنى مشروع جامعة ستانفورد إنشاء ملقمات للتعامل مع بروتوكول نقل النص الفائق، وكذلك للخدمات الأخرى على الويب (بما في ذلك كون تكست Con TEXT، وأداة تلخيص الوثائق في لغة أوراكل Oracle's Document-Summarization Tool).

اللوحة رقم (١٣- ٥)

لجنة العمل الخاصة بالحفظ الأرشيفي للمعلومات الرقمية

شُكّلت لجنة العمل الخاصة بالحفظ الأرشيفي للمعلومات الرقمية في أواخر عام ١٩٩٤م من قبل كل من لجنة الحفظ والوصول، ومجموعة المكتبات الرقمية، بهدف دراسة المشكلات التي تواجه الحفظ الرقمي. وترأس هذه اللجنة جون جاريت John Garrett، الذي عمل فيما بعد في اتحاد مبادرات البحوث الوطنية المعروف بسي إن آر آي CNRI، ودونالد وترز Donald Waters من جامعة ييل Yale University. وكان التقرير الصادر عن هذه اللجنة عام ١٩٩٥م هو أول نظرة شاملة على قضايا الحفظ الرقمي من حيث الجوانب القانونية، والاقتصادية، والتنظيمية، والفنية. وقد ألقى هذا التقرير الضوء على مخاطر التعطل التقني، مع ضرب أمثلة مثيرة على ذلك. فقد لاحظ التقرير - على سبيل المثال - أنه لم تعد هناك برامج آلية تستطيع مثلاً تحليل البيانات التي جمعها مشروع مخزون الموارد الوطنية واستغلال الأراضي في نيويورك the New York Land use and resources عام ١٩٦٠م. ودرست لجنة العمل هذه إمكانية بناء نظام أرشيف وطني يتخذ أساساً للحفظ طويل الأمد. وبالرغم من أن اللجنة أقرت بأهمية - بل بضرورة - هذا النظام، إلا أنها أدركت أيضاً أنه لن يكون حلاً شاملاً لجميع المشكلات، فالمسؤولية الأساسية لحفظ المعلومات تقع على عاتق الأفراد والمؤسسات التي تديرها على حد سواء، وهؤلاء يجب أن يعتبروا أنفسهم جزءاً من النظام الأكبر.

وأكد التقرير على أن الحفظ الأرشيفي للمعلومات هو عملية أكبر من

مجرد نقل المعلومات من وسيلة تخزين متقادمة إلى أخرى، بل يجب أيضاً حفظ وسائل تحويل المعلومات نفسها وترجمتها ومعالجتها. ولهذا السبب من المحتمل أن يصبح نقل المعلومات من شكل إلى آخر ومن نظام آلي إلى آخر هو الأسلوب السائد في عملية الحفظ الرقمي.

ومن اللافت للنظر عناية التقرير بدراسة الجوانب القانونية لعملية الحفظ، حيث إن جانباً كبيراً من المعلومات التي ينبغي حفظها مملوك لهيئات تعدها ملكية فردية سواء كان ذلك لدوافع مالية أو لأية أسباب أخرى. وبطبيعة الحال تتردد هذه الهيئات في توفير نسخ من معلوماتها لحفظها في الأرشيفات. ومن ناحية أخرى، يجب أن تؤكد الأرشيفات على حقها في الحصول على المعلومات للصالح العام، وخاصة المعلومات المعرضة للضياع، الخاصة بإحدى الشركات التي لم يعد لها وجود فعلي في السوق.

إن هذا التقرير يعد دراسة كاملة أعدت في هذا المجال. بالرغم من أن الحسابات المالية الواردة فيه قد أضحت قديمة، هذا فضلاً عن أن التقرير لم يتطرق جدياً إلى تحليل الخيارات والفرص المتاحة.

التخزين :

كان طول عمر المنتج المادي a physical artifact في الماضي يعتمد على طول عمر المادة أو المواد التي صنع منها، وسواءً أكان المنتج مجرد سجلات بسيطة صادرة عن الكنائس والحكومات، أو كنوزاً مثل حجر رشيد، ومطويات [مخطوطات] البحر الميت، وسفر يوم القيامة، وأناجيل جوتنبرج، فإن ما تبقى من هذه المنتجات هي تلك التي صنعت من مادة لا

تتلف بسهولة، وعلى الأخص الورق فائق الجودة.

لا يمكن لأحد أن يضمن بقاء أي من الوسائط الرقمية الحالية لفترة طويلة. بل إن العمر الافتراضي لبعضها، كالشريط الممغنط، يكون قصيراً للغاية، مع أن هناك بعض الوسائط الأخرى، كالأقراص المدمجة، تكون أكثر تحملاً، إلا أنه لا يمكن لأحد التنبؤ بالعمر الافتراضي الأقصى لها. وعلى هذا فإن جميع المعلومات الرقمية ستكون عرضة للضياع في غضون عقود قليلة إن لم يتداركها أحد بالاهتمام. وتصف اللوحة رقم (١٣-٦) بعض الأساليب المستخدمة لتخزين المعلومات الرقمية في الوقت الحالي، مع ملاحظة أن الاهتمام منصب على خفض تكلفة الأجهزة والمعدات وعلى زيادة سرعة استرجاع المعلومات، وليس على إطالة عمر المنتج.

اللوحة رقم (١٣-٦)

تخزين المعلومات الرقمية

تستخدم المكتبات الرقمية أعداداً ضخمة من وسائط التخزين؛ فصفحة واحدة من نص مكتوب وفقاً لرموز المعيار الأمريكي لتبادل المعلومات والمعروف بآسكي ASCII، قد تحتوي على بضعة آلاف من الحروف أو التمثيلات characters، في حين أن صفحة واحدة ملونة ممسوحة ضوئياً لا تتعدى مساحتها بوصة مربعة تحتاج أكثر من ميجابايت (مليون بايت)، كما أن ساعة من التسجيل الصوتي المخزن رقمياً على قرص مدمج تحتاج إلى أكثر من ٦٠٠ميجابايت، بل إن دقيقة واحدة من عرض فيديو قد يحتاج إلى أكثر من جيجابايت من البيانات قبل ضغطها. ولتقليل مساحات التخزين، تُضغط المواد الكبيرة، ويشمل ذلك تقريباً جميع الصور، والتسجيلات

الصوتية، والأفلام. والفكرة الأساسية لضغط الملفات بسيطة مع أن العمليات الرياضية معقدة، فالمعلومات الرقمية تشتمل على فائض من المعلومات الهامشية أو معلومات لا داعي لها redundancy، مثال ذلك أن صورة على صفحة واحدة قد توجد بها مساحات بيضاء غير مستغلة، وعليه فليس من الضروري ترميز كل بكسل فيها على نحو منفصل. وحيث إن الاختلافات بين إطارات الفيديو المتلاحقة طفيفة، فمن الأسهل تسجيل الاختلافات بينها من كل إطار منها على نحو منفصل.

إن الضغط المحكم lossless compression للملفات يقوم بمسح المعلومات المكررة أو الضوضائية على نحو يمكن عكسه تماماً؛ أي يمكن إعادة بناء البيانات الأصلية كما كانت بالضبط، أما الضغط غير المحكم lossy compression فلا يمكن عكسه؛ أي أن التقريب مدعاة إلى فقد بعض المعلومات. وفي بعض التطبيقات يجب أن يكون الضغط محكماً، مثال ذلك في حالة التجارب الفيزيائية، حيث تعد النقطة الواحدة في الصورة ذات أهمية كبيرة، ولذا فإن أي تعديل في الصورة قد يقوّض مصداقية التجربة. وإن كان يمكن في معظم التطبيقات القبول ببعض النواقص. وتعد كل من طريقة JPEG (المستخدمة مع الصور)، و MPEG (التي تُستَخْدم للأفلام) من طرق الضغط غير المحكم والتي تجري معايرتهما لتقديم صورة مريحة للغاية للعين البشرية.

إن طرق ضغط الملفات تصغر حجم البيانات إلى حد بعيد، إلا أن الملفات تبقى كبيرة. فحتى بعد الضغط، نجد أن صفحة أحادية اللون monochrome من نص ممسوح ضوئياً يزيد حجمها على ٥٠,٠٠٠ بايت.

كما أن الضغط بطريقة MPEG يصغّر حجم الفيلم الرقمي من ٢٠ أو ٣٠ ميجابايت في الثانية إلى ١٠ ميجابايت في الدقيقة. وحيث إن المكتبات الرقمية يمكن أن تخزن ملايين من هذه المواد، فإن سعة التخزين تعد من الأمور الهامة للغاية.

إن وسيلة التخزين المثالية للمكتبات الرقمية يجب أن تسمح بتخزين مساحات واسعة من البيانات وبتكلفة قليلة، وأن تكون سريعة في تخزين المعلومات وقراءتها، وأن تكون محل ثقة وأن تدوم طويلاً.

إن الأقراص الممغنطة الدوارة rotating magnetic desk هي وسيلة التخزين القياسية في النظم الآلية الحديثة. وتتفاوت هذه الأقراص من حيث الحجم بين عدد قليل من مئات الملايين من البايتات إلى وحدات تضم آلاف الجيجابايتات. (الجيجابايت يساوي ١٠٠٠مليون بايت). كما أن هذه الأقراص سريعة بما يكفي لمعظم تطبيقات المكتبة الرقمية، حيث يمكن قراءة البيانات من الأقراص أسرع من نقلها عبر الشبكات. وعند قراءة البيانات من القرص يكون هناك تأخر بسيط للغاية (حوالي ٠,١٥ ثانية) وهو الوقت الذي يستغرقه ضبط رؤوس القرص لبدء القراءة؛ ثم تبدأ قراءة البيانات في وحدات كبيرة (حوالي ٢٥ ميجابايت في الثانية في الأوضاع النموذجية). وهذه السمات التي تميز أداء الأقراص تناسب تطبيقات المكتبة الرقمية التي تحتاج لقراءة وحدات ضخمة من البيانات في وقت واحد.

إن انخفاض تكلفة الأقراص يعد إحدى معجزات التقنية؛ فأسعار الأقراص الممغنطة تنخفض تدريجياً أسرع من أسعار أشباه الموصلات، ففي عام ١٩٩٨م كان سعر الأقراص يقدر ببضع مئات من الدولارات لكل

جيجابايت. والتقنية تتقدم سريعاً للغاية بحيث يمكن للمكتبات الرقمية أن تخطط بثقة لعشر سنوات قادمة على أساس أن التكلفة لن تتعدى ٥%، ومن المحتمل جداً أن تصبح أقل من ١% من التكلفة الحالية.

أما نقاط ضعف الأقراص فتتمثل في عدم الثقة في الاعتماد عليها، فالبيانات الموجودة على الأقراص يمكن أن تضيع بسهولة، إما بسبب عطل في أحد المكونات المادية للحاسب، أو بسبب احتمال الكتابة عليها عن طريق برنامج آخر. وللوقاية من عملية الضياع هذه، فمن المعتاد نسخ البيانات بانتظام على وسيلة أخرى، عادة ما تكون شريطاً ممغنطاً. ومن المعروف كذلك وجود بعض المساحات غير المستخدمة على الأقراص، لذلك فمن السهل تصحيح الأخطاء البسيطة آلياً. وفي الواقع لا يمكن الاعتماد على الأقراص أو الشرائط الممغنطة للتخزين طويل الأمد. فالبيانات ترمّز على فيلم ممغنط رقيق يحفظ على سطح ما، وعاجلاً أم آجلاً يتلف هذا الفيلم، وعليه يمكن القول بأن الأقراص تعد وسيلة ممتازة للعمليات الحالية ولكن ليس للحفظ الأرشيفي.

وأحياناً يتم اختزان مجموعات المكتبات الرقمية الكبيرة في مستودعات ذات طابع هرمي. والمستودع النموذجي يتكون من ثلاثة مستويات: الأقراص الممغنطة، والأقراص البصرية، و الشرائط الممغنطة. فالأقراص الممغنطة تكون دائماً متاحة على الخط المباشر online، بحيث يمكن قراءة المعلومات خلال جزء من الثانية. أما الأقراص البصرية فإنها تعد وسيلة أرخص لتخزين كميات كبيرة من البيانات، إلا أن أسطوانات الأقراص تخزن منفصلة offline، وقبل استخدام القرص البصري، يجب أن يقوم الروبوت

ROBOT بنقله آلياً من المستودع silo إلى قارئ الأقراص - وهي عملية تتم ببطء. وكذا فإن الشرائط الممغنطة تخزن أيضاً منفصلة، ويتم حملها بواسطة الروبوت.

وبالنسبة للحاسبات التي تستخدم المخازن الهرمية، يعد المخزن الهرمي نظام ملفات مترابطًا واحدًا. وقد قل اللجوء إلى اتباع أسلوب نقل البيانات من الأقراص الممغنطة الأسرع والأعلى تكلفة إلى وسائط أبطأ وأرخص. ونتيجة للانخفاض المتواصل في تكلفة الأقراص الممغنطة وسعتها، أصبحت الحاجة إلى إيجاد مستوى تخزين متوسط محل تساؤل. فالأقراص والشرائط الممغنطة تؤدي وظائف جيدة وكلاهما ضروري، إلا أن الحاجة إلى مستوى تخزين متوسط قد تنتفي في المستقبل.

إعادة النسخ والتنشيط Replication and Refreshing :

يعد النسخ أحد الأساليب الأساسية لمعالجة البيانات، فالبيانات الهامة الموجودة كنسخة واحدة فقط على حاسب آلي تكون عرضة للمخاطر، فالمكونات المادية للحاسب قد تتعطل؛ والبيانات قد تضيع بسبب مشكلة في البرامج؛ أو قد يقوم موظف غير مؤهل أو غير أمين بمحو البيانات؛ وقد يحترق المبنى الذي يوجد فيه الحاسب، أو يغمره فيضان، أو تصيبه أية كارثة أخرى. لهذه الأسباب، تلجأ مراكز الحاسب الآلي دورياً إلى إعداد نسخ احتياطية من جميع البيانات وتخزنها في مواقع آمنة. أما الهيئات الأكثر تنظيماً فإنها تزيد على ذلك بتجميع السجلات الهامة دورياً لتخزينها لأمد طويل. ومن أساليب الاستنساخ أيضاً الاحتفاظ بنسخ من السجلات المالية والقانونية على شكل مصغر (أي الميكروفيلم أو الميكروفش)، حيث إن

الأشكال المصغرة الأرشيفية الجيدة تدوم أطول من غيرها.

وحيث إن أعمار جميع الوسائط التي تخزن عليها المعلومات الرقمية قصيرة، فيجب أن تخطط المكتبات الرقمية لتنشيط مجموعاتها دورياً على المنوال نفسه. فكل عدة سنوات يجب نقل البيانات إلى وسائط تخزين جديدة. ومن وجهة النظر المالية، لا يشكل هذا الأمر تحدياً كبيراً، إذْ من المتوقع أن يتواصل انخفاض أسعار أجهزة الحاسب الآلي مع زيادة سعتها التخزينية خلال العقود القليلة القادمة. فتكلفة الأجهزة اللازمة لنقل البيانات الحالية ستكون بعد عشر سنوات من الآن نسبة بسيطة من تكلفتها الحالية، ويمكن أيضاً أن تخفف أجهزة الروبوت من ضغوط العمل المطلوبة. وفيما يتعلق بالمكتبات الرقمية، هناك سؤال تنظيمي وهو: هل ستبذل المكتبات والناشرون جهوداً تنظيمية لتنفيذ هذه العمليات ؟

حفظ المحتوى عن طريق النقل Preserving Content by Migration :

حتى لو أخذنا بالافتراض الأكبر المتمثل في تنشيط البتات دورياً من وسيلة إلى أخرى، بحيث يمكن تلبية ضرورة حفظ البيانات الخام أو صيانتها، فإن المشكلات لا تزال في بدايتها. فالمعلومات الرقمية تبقى عديمة الفائدة بدون معرفة ومعالجة قضايا الصيغ والأشكال، والبروتوكولات، وما وراء البيانات. ولا يزال في الإمكان قراءة المخطوطات القديمة، حيث إن اللغات والكتابة تتغير ببطء عبر السنين. وتفسير الوثائق [١] القديمة يتطلب خبرات كبيرة، وإن كانت هذه الخبرات تقل مع مرور السنين. ولا يستطيع

(١) الأفضل تسميتها تحقيق الوثائق (المترجمان).

الباحثون فك رموز الكتابات القديمة إلا بالمثابرة والإلهام.

إن أنماط تطبيقات الحاسب تتغير على نحو مستمر. فقد تصعب قراءة صيغة ملف عمره ١٠ سنوات، كما لا يوجد في العالم حاسب واحد يمكنه تشغيل برامج لبعض الحاسبات الأخرى التي كانت منتشرة منذ فترة قصيرة مضت. وإن كانت بعض هذه الصيغ بسيطة تماماً. فعلى سبيل المثال، إذا عثر أحد الباحثين الأثريين - مستقبلاً وبالمصادفة - على ملف يتضمن نصًّا مكتوبًا برموز آسكي ASCII ضمن بيانات أخرى، فمن المحتمل أن يتمكن من قراءة النص وتفسيره، حتى إذا ضاعت جميع المعارف المتعلقة بصيغ رموز آسكي، حيث إن رموز هذه الصيغة بسيطة للغاية. إلا أن هذا الشكل (ASCII) يبقى حالة استثنائية. فهناك صيغ أخرى شديدة التعقيد ومن الصعوبة بمكان أن نصدق أن أحداً يتمكن من فك رموز صيغة MPEG لضغط الملفات من دون أن يتوافر لـه سجل بالعمليات الرياضية التي بُنيت عليها هذه الصيغة، أو أن شخصاً يمكنه فهم برنامج آلي كبير معتمداً على كود الآلة الخاصة به. لذلك، وعلاوة على تخزين البيانات الخام، يجب أن تحافظ الأرشفة الرقمية على وسائل فهم نوع هذه البيانات وبنيتها والصيغة التي كتب بها، وإذا كان ضرورياً استخدام برنامج آلي لتفسير البيانات فيجب المحافظة على هذا البرنامج وعلى أية وسيلة أخرى يمكنها تنفيذ البرنامج، وإلا فيتعين نقل البيانات إلى أية وسيلة حفظ أخرى.

وعلى المدى القريب، يمكن الاحتفاظ بنظم آلية قديمة لهذا الغرض، إلا أنه يجب الأخذ بعين الاعتبار قصر دورة حياة الحاسبات نفسها. فعاجلاً أو آجلاً،

سيتعطل الحاسب، ولن يتيسر الحصول على قطع الغيار، وأي برنامج يعتمد على هذا الجهاز سيكون عديم القيمة. وعلى هذا يصبح نقل المحتوى ضرورة لازمة.

لقد باتت عملية النقل أسلوباً عملياً قياسياً في معالجة البيانات منذ عقود. فعلى سبيل المثال، تحتفظ صناديق معاشات التقاعد بسجلات للمعاملات المالية لسنوات طويلة. وفي الولايات المتحدة الأمريكية، تحتفظ إدارة الضمان الاجتماعي بسجلات لضرائب الرواتب المدفوعة عن جميع العمال في كل المهن. وتحفظ هذه السجلات على حاسبات، وتغير النظم الآلية دورياً، حيث يجري إحلال الأجهزة الجديدة محل القديمة، كما يجري تنقيح نظم البرمجيات. وعندما تحدث هذه التغييرات، تنقل البيانات من حاسب لآخر، ومن قاعدة بيانات إلى أخرى. والقاعدة الأساسية التي تستند إليها عملية نقل البيانات، هي أن الصيغ التي توضع فيها البيانات، وبنية البيانات هي التي يمكن تغييرها، أما دلالات المحتوى الأساسي فيجب أن تبقى كما هي.

وهناك أسلوب آخر مقترح لحفظ البيانات، وهو "المشابهة أو المحاكاة Emulation" وتتمثل فكرة المحاكاة في تحديد المواصفات التفصيلية لبيئة استخدام الحاسب اللازم لتنفيذ برنامج معين. ومن ثم يتمكن في أي وقت في المستقبل من بناء بيئة محاكية تعمل مثل البيئة الأصلية تماماً. وفي بعض الحالات الخاصة، تعد المحاكاة اقتراحاً معقولاً؛ فعلى سبيل المثال، يمكن تحديد مواصفات برنامج يستحضر صيغة صورة بسيطة، تشبه صيغة JPEG، ومع ذلك تبقى المحاكاة في جميع الحالات الأخرى مجرد وهم. فحتى بيئات

استخدام الحاسبات البسيطة تعد معقدة للغاية بحيث يصعب تحديد مواصفاتها بدقة. فالدمج الدقيق بين عمليات التركيب والدلالات اللفظية، والقواعد الخاصة أمر يتجاوز الفهم، إلا أن الجوانب الدقيقة الخفية لأي نظام تبقى عنصراً حاسماً للتنفيذ السليم.

الآثار الرقمية القديمة [علم الآثار الرقمي] Digital Archeology :

تمر المجتمعات بفترات عصيبة أثناء فترات الركود والحروب والاضطرابات السياسية، وفي هذه الفترات يصبح نقل المواد الأرشيفية في أدنى سلم الأوليات. ويمكن أن تبقى المنتجات المادية منسية لقرون طويلة في دهاليز الأقبية والمخازن حتى يعاد اكتشافها. أما المعلومات الرقمية فهي أقل تحملاً. وتتناول اللوحة رقم (١٣-٧) كيف صاحب إحدى الفترات العصيبة - وهي انهيار ألمانيا الشرقية - فقد أرشيف الدولة. ويطلق على عملية استرجاع المعلومات من مصادر البيانات التالفة غير المكتملة والقديمة، اسم الآثار الرقمية القديمة أو بالأحرى علم الآثار الرقمي.

بناء مكتبات رقمية بالأرشفة الذهنية Archiving in Mind :

بما أن عملية الأرشفة تنطوي على مخاطر كثيرة، فما الذي يمكننا القيام به الآن لتعزيز احتمال تمكن علماء الآثار القديمة الرقميين من تفسير البتات التي اخترعناها؟. هناك بعض الخطوات البسيطة التي من المحتمل أن تشكل فارقاً كبيراً. تتمثل الخطوة الأولى في تخزين المعلومات في صيغ يتبنى استخدامها على نطاق واسع وفي الوقت الحاضر. وهذا سيزيد من فرصة توافر برامج نقل إلى صيغ جديدة في حالة توقف استخدام الصيغ الأولى. فعلى

سبيل المثال فإن صيغتي لغة ترميز النص الفائق HTML، وبي دي إف pdf، تستخدمان على نطاق واسع في عالم الصناعة مما يجعلنا متأكدين من توافر مستخدمين لهما في المستقبل.

ومن الاقتراحات الجديرة بالاهتمام بناء أرشيف يشتمل على تعريفات الصيغ، ومعايير ما وراء البيانات القياسية والبروتوكولات، والدعائم الأخرى للمكتبات الرقمية. وهذا الأرشيف يجب أن يحفظ على وسيلة حفظ دائمة مثل الورق أو الميكروفيلم، ويجب أيضاً وصف كل شيء بنصوص بسيطة، وفي حالة الحفاظ على الصيغ وأساليب الترميز، فسيمكن فك رموز معظم المعلومات. وقد يعاني علماء الآثار الرقميون في المستقبل من مشقة إيجاد مترجم يمكنه حل الصيغ أو مجموعات التعليمات المهجورة منذ زمن طويل، وإن كان هذا ليس أمراً مستحيلاً. إن صيغ استخدام الحواسيب الحديثة معقدة. ومع أن العالم الأثري الرقمي قد يتمكن من إعادة هندسة البنية الكاملة لأحد الحاسبات القديمة من إنتاج شركة IBM عن طريق استخراج محتويات الذاكرة، فإنه سيقف عاجزاً أمام المواد الأكثر تعقيداً في حالة عدم توافر المواصفات الأساسية للنظام.

وقد يكون "الانتقاء selection"، هو أهم الوسائل التي تستطيع بها المكتبات الرقمية دعم عملية الأرشفة. ولكن ليس مطلوباً ولا ضرورياً الاحتفاظ بكل شيء. فمعظم المعلومات معدة لتكون دورة حياتها قصيرة، ومعظم المعلومات إما أن تكون ثانوية أو عديمة القيمة. كما إن الناشرين يتخذون قراراتهم دائماً بما ينشرون وما يرفضون، وحتى أكبر المكتبات لا

تحصل إلا على نزر يسير من الإنتاج الفكري العالمي. إن المكتبات الرقمية ليست إلا مجموعات معلومات خاضعة لإدارة معينة. وأحد العناصر الأساسية لهذا النظام الإداري هو اتخاذ القرار بشأن ما ينبغي الحصول عليه، وما ينبغي تخزينه، وما ينبغي الحفاظ عليه للمستقبل، وما ينبغي تجاهله واستبعاده.

اللوحة رقم (١٣– ٧)

علم الآثار الرقمي في ألمانيا

في عام ١٩٩٨م، ظهر مقال في صحيفة نيويورك تايم يتناول التحديات التي يواجهها موظفو الأرشيف في العالم الرقمي في حالة عدم حفظ البيانات على نحو متواصل منذ أول وجود لها.

وفي عام ١٩٨٩م، عندما انهار سور برلين وتوحدت ألمانيا مرة أخرى، كانت السجلات أو الوثائق الرقمية لألمانيا الشرقية في حالة فوضى عارمة، وحصل الأرشيف الاتحادي لألمانيا على مجموعات ضخمة من البطاقات المثقوبة، والأقراص الممغنطة، وأشرطة الحاسب الآلي التي تحتوي على وثائق ألمانيا الشرقية سابقاً. وكانت معظم هذه الوسائط في حالة يرثى لها، وكانت البيانات محفوظة في نماذج غير موثقة، وكانت مراكز الحاسب الآلي التي توجد بها هذه المواد قد أغلقت أو خُصخصت على استعجال. ومنذ ذلك الحين، يحاول فريق من الأرشيفيين الألمان إعادة بناء وثائق ألمانيا الشرقية وسجلاتها، وقد أطلقوا على أنفسهم اسم "علماء الآثار الرقميون Digital

archeologists".

وكانت المشكلة الأولى التي واجهت هؤلاء الموظفين هي استرجاع البيانات من الوسائل المخزنة فيها. أما البيانات المخزنة حتى على أفضل الشرائط الممغنطة فكان عمرها الافتراضي قصيراً، حيث إن الشرائط نفسها كانت في حالة سيئة إلى درجة أن هؤلاء الموظفين لم يتمكنوا من قراءتها سوى مرة واحدة فقط. وفي حالات كثيرة، كانت البيانات مخزنة في نظم روسية Russian system لا تدعمها الحاسبات الآلية الأخرى. وعلى الرغم من أن هؤلاء الموظفين تمكنوا من الحصول على العديد من أجهزة الحاسبات الآلية روسية الصنع، إلا أنهم لم يتمكنوا من قراءة حوالي ٣٠% من البيانات المخزنة.

وعندما نسخت البيانات على وسائل أخرى، كانت المشكلات لا تزال مستعصية الحل، حيث كان هناك جزء كبير من البيانات مضغوطاً بأساليب مبهمة وغير موثقة لغرض توفير المساحة. وقد أثارت إحدى قواعد البيانات الهامة للموظفين الشيوعيين بعض هذه المشكلات. وحيث أن برامج قاعدة البيانات الحاسب الآلية التي حُمِلتْ عليها هذه البيانات كانت مستنسخة أصلاً من منتجات شركة IBM، فلم يكن استعادة قاعدة البيانات نفسها أمراً صعباً، إلا أن تفسير البيانات في ظل عدم وجود أي توثيق لها كان أمراً شاقاً للغاية. وكانت هناك ميزة واحدة لصالح موظفي الأرشيف، وهي أنه كان بإمكانهم إجراء مقابلات مع بعض الأفراد الذين قاموا بإنشاء قواعد البيانات هذه، واستفادوا من الخبراء من بين هؤلاء الأفراد في تفسير جانب كبير من

المعلومات وفي حفظ هذه المعلومات وصيانتها والمحافظة عليها.

وقد لخص مايكل فيتنجيت Michael Wettenget كبير الأرشيفيين الألمان، الموقف بوضوح قائلاً: "إن تقنية الحاسب الآلي وجِدتْ من أجل معالجة المعلومات، لا من أجل اختزانها لآماد بعيدة".

الفصل الرابع عشر

المكتبات الرقمية والنشر الإلكتروني في الوقت الحاضر

المكتبات الرقمية والنشر الإلكتروني حقائق واقعة وموجودة في حاضرنا، وليستا فكرة أكاديمية قيد نقاش أو مجرد حلم خيالي، وكنت أتطلع في هذا الفصل الأخير إلى التنبؤ بما هو منتظر، إلا أن تاريخ استخدام الحاسبات يكشف لنا إلى أي مدى يمكن أن تكون مثل هذه التنبؤات عقيمة، وعليه فقد رأيت أنه من المفيد الاحتفاء بالحاضر لا التنبؤ بالمستقبل.

ولا يعني هذا أن نتجاهل المستقبل، فما يجري الآن من أعمال ستصبح إنتاجاً في القريب العاجل، كما أن كل المشروعات الجديدة لا تدخل في إطار التوجه العام، لكن فهم ما يجري الآن سيساعدنا على التبصر بما ينتظرنا مستقبلاً، وتجدر الإشارة إلى أن تأليف هذا الكتاب قد استغرق ما يزيد على اثني عشر شهراً ما بين عامي ١٩٩٧م، ١٩٩٨م، وإن كانت هذه الفترة قصيرة إلا أن المكتبات الرقمية قد حدث فيها تطورات سريعة. ويسعى هذا الفصل إلى عرض بعض الاتجاهات الجارية، لا بهدف التنبؤ بما سيكون عليه الحال في المستقبل، ولكن لفهم ما يجري في الحاضر.

ولعل أحد أسباب صعوبة التنبؤات يتمثل في أن التقنية في الوقت الحاضر تتميز بأنها أكثر نضجاً من الاستخدامات التي توظف فيها، أما في الماضي فقد أثبتت التنبؤات الخاصة بالتقنية الأساسية أنها كانت دقيقة إلى

حد كبير، وفي كل عام تصبح أشباه الموصلات[١] والأجهزة الممغنطة Magnetic devices أصغر وأرخص وأسرع بجانب تزايد قدراتها التخزينية، وهناك أسباب هندسية طيبة تشير إلى استمرار التنبؤ باستمرار هذه الاتجاهات على مدى السنوات الخمس أو العشر القادمة. كما أن هناك بالفعل التزامات بتخصيص الاعتمادات المالية اللازمة لشبكات الاتصال فائقة السرعة، والتي ستتاح في غضون السنوات القليلة القادمة كذلك، لكن التنبؤات الخاصة بالتطبيقات الجديدة التي سيتم تطويرها كانت دائماً توقعات غير دقيقة بالقدر الكافي، حيث ظهرت مثل تلك التطبيقات الأولية، كما هو الحال في الجداول الإلكترونية والنشر المكتبي ومتصفحات الويب فجأة ودون سابق إنذار. ومن ثم فإنه حتى لو لم تحدث قفزات مشابهة، فإن التنبؤات الخاصة بكيفية استخدام التطبيقات والأثر الاجتماعي للتقنيات الجديدة لا يمكن التعويل عليها، والقاعدة الوحيدة التي يمكن التعويل عليها في هذا الصدد تتمثل في أنه عندما يبدأ العالم بالقول : "من الأمور الحتمية أن..."؛ فاعلم أن ما سيصرح به بعد ذلك سيكون حتماً خطأ.

ومن التفسيرات المحتملة للوضع الحالي أن المكتبات الرقمية تعيش في نهاية مرحلة تمهيدية، وعلى وشك أن تبدأ مرحلة جديدة، ويمكن أن نعتبر هذه المرحلة الأولى بمثابة نقلة من عالم المطبوعات التقليدية والمجموعات

(١) مواد توصل التيار الكهربائي بدرجة أقل من المعادن وأكبر من المواد العازلة مثل الزجاج أو الخزف، أي أن مقاومتها تقع بين مقاومة الموصلات ومقاومة العوازل، ومن أمثلة تلك المواد: الجرمانيوم، والسليكون والسلنيوم (المترجمان).

المكتبية إلى الشبكات الرقمية. وتندرج تحت هذه الفئة كل الصحف المتاحة على الإنترنت، والإصدارات الإلكترونية من الدوريات العلمية، والمواد التاريخية التي تم تحويلها إلى الأشكال الرقمية، وبشكل أساسي تسعى كل هذه المواد إلى استخدام التقنية الجديدة لتعزيز أنواع مستقرة من المعلومات. وإذا كان هذا التصور صحيحاً، فإن المرحلة التالية ستشهد بزوغ أنواع أخرى من المجموعات والخدمات التي ليس لها نظائرها التقليدية، أما الأشكال التي سوف تتخذها تلك الأنواع، فلا شك أنه يكاد يكون من المستحيل توقعها الآن.

نظرة قاصرة للمكتبات الرقمية :

حينما يسترجع المرء النظر في أي فترة ماضية، تبدو له الاتجاهات والأحداث الرئيسية واضحة على نحو لم تكن عليه آنذاك، وأمثالنا ممن يعملون في مجال المكتبات الرقمية لا يتمتعون إلا بنظرة قاصرة عن هذا المجال، غير أن هناك بعض المرئيات الخاصة للكيفية التي يبدو عليها، وخاصة لمن تيسرت له أسباب الاطلاع على بواطن الأمور.

ولنا هنا أن نستعير تشبيهاً مفيداً لهذا الموقف، وهو التناقض القائم بين قارب صغير وحاوية ضخمة، فبمقدور القارب الصغير أن يجري ويغير اتجاهاته بسرعة لكن دون أن يخلّف ذلك زخماً يذكر، وكانت المكتبات الرقمية في أول عهدها كما هو الحال في مشروع ميركوري بجامعة كارنيجي ميلون أشبه بالقوارب الصغيرة؛ إذ حققت تقدماً سريعاً، ولكن عندما خبت جذوة الحماس أو نفذ التمويل، فقدت زخمها وقوتها. أما المكتبات التقليدية ودور النشر وغيرهما من المؤسسات التجارية الراسخة فهي أشبه بالحاويات العملاقة، فهي تتحرك بروية، وتغير اتجاهاتها ببطء، ولكن ما إن

تقصد اتجاهها الجديد فإنها تنطلق فيه على نحو منتظم. وفي مجال المكتبات الرقمية والنشر الإلكتروني، يتطلب النجاح الاهتمام بآلاف التفاصيل التي تحول الأفكار الجيدة إلى خدمات عملية. وفي ظل ما تعيشه الإنترنت والويب من مظاهر النضج، بدأت كثير من المنظمات تتجه إلى الاستثمارات طويلة الأمد في المجموعات المكتبية وفي المطبوعات الإلكترونية والخدمات التي تقدم على الخط المباشر عبر الإنترنت. وعليه يمكن القول بأن الحاويات العملاقة في طريقها إلى تغيير اتجاهها.

وعلى مدار عامي ١٩٩٧م و ١٩٩٨م، نضجت كثير من التطورات التي كانت في طور الإعداد على مدى عدة سنوات، ووصلت إصدارات الصحف الإلكترونية إلى درجات عالية من الجودة، بل بدأ بعضها ينافس الإصدارات المطبوعة. كما أصبحت المطبوعات العلمية الهامة، سواء ما يصدر عن ناشرين تجاريين أو ما يصدر عن ناشرين حكوميين، متاحة عبر الإنترنت، كما كانت هذه المرحلة كذلك بمثابة فترة هامة في تاريخ مشروعات التمويل التي تبنتها بعض المكتبات مثل مشروع جستور ومكتبة الكونجرس، حيث تزايدت أعداد المواد المتاحة في سياق مثل هذه المشروعات على نحو كبير، كذلك حقق مشروعا "دبلن كور" لما وراء البيانات، استخدام معرفات الكائنات الرقمية للمطبوعات الإلكترونية دفعات قوية، وبدأ وكأنه يحقق تحولاً من وضعه الهش إلى مكانته القوية السائدة الآن.

وعلى الجبهة التقنية، غزت المنتجات المخصصة لأغراض الذاكرة التلقائية المطابقة أو الخفية Mirroring and caching السوق، وتوافرت وسائل أمن الويب، وأخيراً استخدمت لغة جافا للبرمجة على نطاق واسع. ومن

الملاحظ أن كل هذه التطورات بدأت منذ سنوات مبكرة، ولا يمكن اعتبار أي منها مجرد مجال للبحث، وعلى وجه الإجمال يمكن القول بأن كل هذه التطورات تعكس أوجه التقدم الهائل.

وفي الولايات المتحدة الأمريكية، نمت التجارة الإلكترونية عبر الإنترنت بسرعة، كما أصبح من المألوف وعلى أوسع نطاق، شراء الكتب وتذاكر الطيران والأسهم والسيارات عبر الإنترنت، ومن جهة أخرى تحرص إدارة الموارد الداخلية على حث الناس على تسديد ضرائب الدخل من الإنترنت، كما أقر الكونجرس الأمريكي تعديلات ومراجعات هامة على قانون حق المؤلف، ونظراً لتوافر الأموال المتاحة للاستثمار وانفتاح الأسواق أمام المنتجات الجديدة، أصبحت أسهم الإنترنت مجالاً مفضلاً للمتضاربين في سوق الأسهم، كذلك تشهد السوق انتعاشاً كبيراً في سوق بيع الحاسبات الآلية الشخصية التي تقل قيمة الواحد منها عن ألف دولار، مما كان له أثره الواضح على إتاحة المعلومات المباشرة عبر الإنترنت ووصولها إلى قطاع عريض من الناس.

وعلى الجانب الآخر، يمكن القول بأن ما من عام يمر إلا ويحدث ما يجرح كماله، ويكون بوسع المتشائمين ما يرجون له من الأحداث المزعجة، فخلال عام ١٩٩٧م وصل عدد الرسائل غير الأخلاقية المرسلة عبر البريد الإلكتروني إلى معدلات مثيرة للضجر، كما أصدر كبار منتجي البرمجيات نسخاً غير متوافقة مع لغة الجافا، واستمرت سياسة الولايات المتحدة الخاصة بفك التشفير تسير بشكل يسير أشبه بالنعامة، وكانت كل هذه من المشكلات قصيرة الأجل. وبمثول هذا الكتاب للنشر، آمل أن تتم السيطرة

على رسائل البريد الإلكتروني غير الأخلاقية، وأن يصدر معيار خاص بلغة الجافا، وهناك على أية حال بصيص من الأمل بشأن سياسة الولايات المتحدة نحو فك التشفير Encryption.

وعلى مدى السنوات القليلة القادمة، يمكن توقع حدوث تطورات تدريجية مشابهة لتلك التي حدثت في عامي ١٩٩٧، ١٩٩٨م، ولكن يمكن تلخيص توجهها على نحو بارع في عبارة مفادها أن أعداداً كبيرة من المتحمسين للإنترنت أو النشطاء تجاه خدماتها يترقبون الفرص التي توفرها الإنترنت لتقديم منتجات وخدمات جديدة.

الناس :

من وجهة نظر قاصرة، إذا كان من السهولة لنا أن نتعرف إلى الأنشطة الفردية، فإن من الصعوبة فهم التغيرات الجوهرية التي تعد جزءاً من طبيعة هذه الأنشطة؛ فالاتجاه الذي يعتبر أمراً جوهرياً على المدى الطويل، قد يكون من أصعب الأشياء التي يمكن قياسها على المدى القصير، فكيف إذن بتغير عادات الناس؟ فكثير من هؤلاء الناس يكتبون وينشرون على صفحات الويب ويقوم مصممو الرسومات بإتاحة أعمالهم على صفحات الويب كذلك، وفي المقابل يقوم آخرون بقراءة ومشاهدة هذه المواد، فمن هؤلاء الناس؟ وماذا كانوا سيفعلون بدلاً من ذلك ؟

حقاً لقد تغيرت عادات الناس تغيراً كاملاً، قم مثلاً بزيارة إحدى المدن الجامعية، أو أي مؤسسة أخرى تحرص على الإفادة من المعلومات على نحو مكثف، فسوف يتضح لك أن الناس يقضون الساعات الطوال على مدار الأسبوع أمام شاشات الحاسبات، مستفيدين من المعلومات المتاحة عبر

الإنترنت الآن، وفي المنزل هناك مؤشرات على أن قسطًا كبيراً من الوقت الذي يقضيه الناس في التعامل مع الإنترنت الآن كانوا يقضونه فيما سبق في مشاهدة التلفاز، وفي العمل هناك تساؤل: هل يقرأ الناس أكثر، وهل استبدلوا الأنشطة التقليدية كزيارة المكتبة بالمعلومات التي يحصلون عليها من الإنترنت ؟ إننا حينما نحاول أن نسترجع فترة التغير الحالية بعد عشر سنوات من الآن، ربما ستصبح هذه الاتجاهات واضحة لنا، ومع ذلك لا نملك الآن بحكم ما يتاح لنا من معلومات شحيحة، أن نفترض أو أن نستقرئ بشكل مفرط، وإليك بعض التخمينات المبنية على ملاحظة شخصية وعلى إحساس داخلي قوي.

إن الإثارة التي تتمتع بها المعلومات المباشرة المتاحة على الإنترنت جذبت أعداداً كبيرة من الناس لمجال المكتبات الرقمية، كما أن الذين كانوا ينظرون إلى مهنة المكتبات على أنها مهنة مملة، وإلى النشر على أنه حكر على الكتب دون غيرها، أصبحوا مفتونين بإنشاء وتصميم المواد وإتاحتها على الإنترنت، وهذا الحماس وتلك الطاقة اللذان قام بتوفيرهما نشاط هؤلاء الرواد الجدد للإنترنت يؤثران على المهن القديمة بشكل جوهري وبقدر أكبر مما يتوقع أي شخص، وعلى الرغم من أن لكل مجموعة رؤيتها، فإن الكثيرين يجدون متعة بالغة في الفرص الجديدة المتوافرة لهم.

وعندما بدأت الويب عملها، لم تكن الخبرة بالإنترنت متوافرة للجميع لدرجة أن أي شخص كان لـه مهارة متواضعة في هذا المجال كان بمقدوره أن يتقاضى راتباً مرتفعاً، والآن وعلى الرغم من أنه لا يزال هناك طلب شديد على الخبراء الحقيقيين، فإن مستوى المهارة أصبح عالياً بشكل عام.

ويتمثل أحد دلائل هذا التغير في نمو البرامج التي تساعد من هم في مقتبل العمر على تعلم المجالات الجديدة. وفي الولايات المتحدة تقدم كل كلية من كليات المجتمع دورات في الإنترنت والويب، كما تزايدت بشكل هائل الشركات التي تقدم البرامج التدريبية في الحاسب الآلي، وفي السياق نفسه شهدت البرامج المتخصصة في المكتبات الرقمية تضخماً في عدد المشتركين فيها.

وفي عام ١٩٩٧م طُلب من أحد طلاب جامعة كورنيل أن يبحث عن بعض المعلومات في مكتبة الجامعة، فسأل قائلاً: "من فضلك هل بمقدوري أن أستخدم الويب ؟ فإنني لا أجيد التعامل مع المكتبات التقليدية"، ومنذ فترة قريبة، قالت إحدى أعضاء هيئة التدريس بجامعة كاليفورنيا في بيركلي مازحة: إن كلمة "المكتبات الرقمية" قد استُنزفت، ولم يعد لها معنى من كثرة التكرار، حيث ترى أن الإنترنت بالنسبة للطلبة هي المكتبة. فهل هؤلاء الطلاب يعتقدون أن مكتبات بيركلي التقليدية الجميلة ستكون في المستقبل بديلاً ملموساً لشيء حقيقي؟ وهل هذه هي آراء متطرفة لأقلية متميزة؟ أم هل هي آراء نافذة البصيرة للجيل القادم من مستخدمي المكتبات ؟ لا أحد يعرف ذلك؛ لأن البيانات متناثرة ومتضاربة، ولدينا بعض النتائج الإحصائية التي تم انتقاؤها من عدد من المصادر، ونوردها في الفقرات التالية راجين أن تؤخذ بنوع من الحذر.

ففي مسح أجري في بتسبيرج عام ١٩٩٦م، تبين منه أن ٥٦% ممن تتراوح أعمارهم بين ١٨-٢٤ سنة استخدموا الإنترنت في مقابل استخدامها من قبل ٧% ممن هم فوق سن الخامسة والخمسين. وفي استطلاع للرأي

أجري على مستوى الولايات المتحدة عام ١٩٩٧م، تبين أن ٦١% ممن هم في سن المراهقة (وتحديداً ٦٦% من البنين، ٥٦% من البنات) استخدموا الويب. وفي عام ١٩٩٦م كذلك تبين من إحدى الدراسات أن ٧٢% من الأطفال الذين تتراوح أعمارهم بين ٨ - ١٢ سنة كانوا يمضون بعض الوقت أمام شاشات الحاسبات خلال شهر محدد.

وفيما يتصل بالشبكة التي بدأ كثير من الشباب ينخرطون في استخدامها كمكتبة، أشير إلى أنها كانت تضم ما يقرب من ٥,٣ مليون حاسب آلي في عام ١٩٩٨م، وقدرت دراسة أخرى دقيقة في أواخر عام ١٩٩٦م أن عدد مواقع الإنترنت التي باتت شائعة الاستخدام من قبل الجميع تقريباً تقدر بمليون موقع، وتتاح هذه المواقع على ٤٥٠ ألف حاسب مضيف، كما أن ما يقرب من ٣٠٠ ألف من هذه المواقع هي من المواقع المستقرة. كما أن عدد الصفحات التي كانت مهيئة بلغة تهيئة النصوص الفائقة ومتاحة على خوادم عامة قد وصلت إلى ما يقرب من ٣٢٠ مليون صفحة. وأيًا كان العدد بالضبط، فلا يختلف اثنان على أنه عدد ضخم، وأنه ينمو بشكل متسارع.

وربما إذا ما قرأ أحد المتشائمين هذه الأرقام فسيفسرها بأن الشباب في الولايات المتحدة قد آمنوا بأهمية المعلومات المتاحة على الويب، كما أن من هم في مقتبل أعمارهم يجتهدون لتقبل هذه الوسيلة، أما كبار السن الذين يرسمون الخطط ويتحكمون في الموارد، فقد تجاهلوا استخدام هذه الوسيلة. ومع ذلك تظهر الملاحظة أن مثل هذا التحليل يرسم صورة كئيبة على نحو غير ملائم، أما الحقيقة التي تتجسد في الاستثمار القوي لكثير من الهيئات الكبرى في مجال المعلومات الرقمية تظهر أن بعض القادة القدامى على

الأقل يؤمنون [بتقنيات] العالم الجديد.

الهيئات :

على الرغم من أن كثيراً من المنظمات حريصة على الاستثمار في مجال المكتبات الرقمية ومصادر المعلومات الإلكترونية، فإنه ليس بمقدور أي أحد أن يعرف بشكل مؤكد نوعية الهيئات التي من المحتمل أن تحرز نجاحاً كبيراً في هذا الصدد. ففي بعض الأحيان، قد يكون حجم المنظمة من عوامل النجاح، ومع ذلك فهناك منظمات صغيرة وذكية تحقق نمواً وازدهاراً في هذا المجال.

وفي عام ١٩٩٥م نشر مقال في مجلة "الإيكونومست The Economist" وصف عملية التحكم في الإنترنت بأنها أشبه بعلاقة بين العمالقة والنمل، فالعمالقة هم الشركات الضخمة كشركات الهواتف، وعمالقة وسائل الإعلام، أما النمل فهم الأفراد الذين لم يكن لهم من أسباب القوة عندما كان ينظر إليهم فرادى غير أنهم نجحوا في تشكيل الإنترنت إجمالاً، وغالباً ما كان هذا النجاح مناقضاً على نحو مباشر للمصالح المعلنة للعمالقة.

وبشكل خاص نجح مجتمع النمل في الإبقاء على الإنترنت وما يصحبها من عمليات كمجال مفتوح، وفي خلال السنوات القليلة الماضية شهد مجالا المكتبات والنشر الإلكتروني آفاق الاندماج في الوحدات التنظيمية الضخمة، وتمثل ذلك في عالم المكتبات الرقمية في حركة ما يعرف بالاتحادات، في حين تمثل في مجال النشر [الإلكتروني] في اندماج الشركات الضخمة. ومع ذلك ورغم تشكل هذه الكائنات العملاقة فقد استمرت قوة النمل أو الأفراد، وقد بدا في الوقت الحاضر أن العمالقة والنمل يمكن أن يتعايشا ويزدهرا،

وفي هذه الأثناء أصبح بمقدور من كانوا ينتمون إلى مجتمع النمل، مثل محرك البحث ياهو وشركة أمازون، أن يكونوا من العمالقة الجدد.

المجموعات ومدى إتاحتها أو الوصول إليها :

ثمة عاملان سيؤثران بشكل كبير على مستقبل المكتبات الرقمية، وهما: المدى الذي ستتاح عنده المجموعات جيدة التنظيم على الإنترنت، ونماذج العمل التجاري الذي ستنشأ في هذه البيئة.

لقد تزايدت أعداد المواد المتاحة على الإنترنت بمعدلات كبيرة جداً، وتظهر معدلات اتساع الويب أنه لا سبيل للتراجع أو التباطؤ، كما تزايدت أعداد المواقع الجيدة بشكل كبير، وتعد المواقع التي تشغّلها المؤسسات الصحفية ووكالات الأنباء من الأمثلة الطيبة للنماذج التي يحتذى بها في هذا الصدد، مع ملاحظة أن هذه المواقع تُظهر مدى قابلية المعلومات الرقمية المتاحة للاختراق. كما أن خدمات الأخبار الإلكترونية كتلك التي تقدمها صحف سيدناي مورننغ هيرالد Sydney morning Herald والنيويورك تايمز New York Timesوالسي إن إن CNN تعمل على تقديم أحدث الأخبار وبطريقة عرض جيدة للغاية للمستخدمين ودون دفع أية رسوم. ومن المحتمل أن تتجاوز معدلات قراءة أي من هذه الصحف معدلات قراءة أي صحيفة أمريكية، وهي مصادر معلومات ممتازة غير أنها تتسم بسرعة زوالها؛ حيث تتغير المعلومات بصورة مستمرة، ففي نهاية اليوم تختفي معظم تلك المعلومات. وإذا كانت المكتبات التقليدية تحرص على تجميع الصحف واختزانها لقرون لاحقة، فليس ثمة مكتبة أو أرشيف يقوم باختزان مثل هذه المواقع الإلكترونية.

ومن الصعوبة تحديد جميع التخصصات العلمية التي تمتلك أكبر قدر من المواد المكتبية القياسية المتاحة عبر المكتبات الرقمية، ذلك لأن نسبة ما يتاح الآن من إنتاج فكري علمي وتقني عبر هذه المكتبات لا تزال محدودة، وإنما النسبة الكبيرة منه تشغلها المعلومات الحكومية. وإذا كانت المعلومات القانونية تتوافر منذ فترات طويلة على الإنترنت، وذلك مقابل رسوم باهظة جداً، فإن المعلومات ذات الصبغة التجارية والطبية تتاح على نحو متنوع ومتفاوت. وتعمل المكتبات العامة على توفير إمكانية الوصول لمصادر المعلومات عن الأحداث الجارية كالصحف وجداول مواعيد رحلات الطيران، والإعلانات عن الوظائف، ونماذج تسديد الضرائب وإتاحة هذه المصادر على الإنترنت عادة بالمجان.

ويتوافر في كثير من المجموعات معلومات حديثة متاحة في أشكال رقمية مع أنها ليست معلومات تاريخية، هذا بالرغم من وجود مشروعات مزدهرة لتحويل المواد التقليدية إلى أشكال رقمية وإتاحتها في المكتبات الرقمية؛ فالمكتبات تعمل على تحويل مجموعاتها التاريخية، كما يسعى الناشرون إلى ذلك بتحويل الإصدارات القديمة من الصحف والجرائد، وتعتزم المشروعات القائمة زيادة معدلات عمليات التحويل، كما أن هناك مشروعات جديدة في طريقها للنمو. ولقد قامت بالفعل مشروعات كثيرة بتحويل أكثر من مليون صفحة، وكان تدشين أولى خطط التحويل لبليون صفحة قد تم في عام ١٩٩٨م.

وربما تكون أكبر هوة في مجال التحويل إلى الأشكال الرقمية تلك

الموجودة في مجال الأعمال الترفيهية الذي يتخذ بعداً تجارياً، فقد خاض بعض منتجي وسائل الترفيه، كالأفلام السينمائية والبرامج التليفزيونية والإذاعية والروايات والمجلات، تجارب استخدام الإنترنت في هذا الصدد، ولم يكن لها سوى الأثر القليل. وربما يعود ذلك في جانب منه إلى المحدوديات التقنية المتصلة بالإنترنت في هذا الجانب، حيث تعد الصور المنقولة عن طريق الكابلات التليفزيونية، والتي يتلقاها المشاهدون أفضل بكثير من تلك التي تبث عبر الشبكة أو التي يتم نقلها إلى الحاسبات الآلية الشخصية، معقولة الثمن. وإلى حد ما يمكن القول بأن معدل التغيير هو الذي تمليه الممارسات التجارية، ومع أن الترفيه يدخل في إطار الأنشطة التجارية الكبيرة، فإن صناعة هذا النشاط لم تستطع أن تكتشف بعد إلى أي حد يمكن استخدام الإنترنت على نحو مربح في أنشطتها.

إن إتاحة الاتصال المفتوح هو السمة الدائمة للمكتبات الرقمية، غير أن استخدام النماذج التجارية الجيدة في هذا المجال لم يستكشفه بعد إلا النزر اليسير من الخدمات، وقليلة هي مواقع الإنترنت التي استطاعت أن تحقق أرباحاً مالية جيدة عن طريق الإعلانات التي تنشر عبر صفحاتها، بل إن معظم هذه المواقع لا تزال تتلقى دعماً مالياً خارجياً. إن طبيعة العمل بالمكتبات الرقمية وصناعة النشر الإلكتروني تحتاج إلى توافر متخصصين محترفين مهرة، يعملون على إنتاج المعلومات وتنظيمها وإدارتها، وهي كوادر مكلفة على أية حال. وفي النهاية يمكن القول بأن نوعاً ما من التوازن الاقتصادي سيظهر في الأفق وستقوم بمقتضاه بعض المجموعات بإتاحة

الوصول المجاني للمستفيدين منها، بينما ستقوم مجموعات أخرى بإتاحة هذا الاتصال نظير رسوم مباشرة، ولم يحن الوقت بعد لاستكشاف الوضع الذي سيكون عليه هذا التوازن.

التقنية :

إن تقنية الويب التي وهجت شعلة النمو الأخير تزدهر يوماً بعد يوم، فخلال منتصف التسعينات تطورت الويب بسرعة مذهلة مما دفع الناس إلى صياغة مصطلح "عام الويب" ليصفوا به فترة قصيرة من الزمن مشحونة بالكثير من التغيرات التي بدت وكأنها سنة شمسية كاملة. ومع ازدهار الويب تباطأت خطى التغيير في التقنية لتصل إلى معدلها الطبيعي في مجال استخدام الحاسبات، وكل عام يحمل بين طياته نوعاً من التغير التدريجي، وعلى مدى سنوات قليلة أصبحت التغيرات التدريجية هذه أمراً أساسياً، غير أن خطى المجموعة تباطأت. ولكن هذا لا يعني أن حركة النمو التي تشهدها قد وصلت إلى نهايتها، فهي أبعد ما تكون عن هذه المرحلة، فقد تزايدت أعداد المواقع على الويب بشكل متسارع، بحيث تشير إحصاءات المواقع المزدحمة بأن حجم الاستخدام يتضاعف سنوياً، وعلى جبهة أخرى ترتقي جودة الأعمال التصويرية وغيرها من معايير الخدمات على نحو منتظم. كما تحمل التقنيات العديدة الصاعدة بين طياتها آمالاً واعدة للمكتبات الرقمية والنشر الإلكتروني، ومن بين هذه التقنيات ما يلي: الأسماء الثابتة مثل محددات الكائنات الرقمية، والتهيئة عن طريق لغة التهيئة الموسعة، وإطار وصف المصادر، ونظام اليوينكود للترميز. ولاشك أن نجاح هذه التقنيات

سيعتمد على تقلبات السوق؛ حيث سيؤدي التقبل الواسع لأي من هذه التقنيات أو جميعها إلى ثمرات كثيرة ستكون من صالح المكتبات الرقمية.

إن أداء الإنترنت مستمر في التحسن على نحو متميز، وقد أثمرت السنوات القليلة الماضية عن سلسلة من المبادرات الحكومية والتجارية التي تهدف إلى تحقيق قفزات كبيرة في خطى هذا الأداء وفي مدى موثوقية ما يتاح عبر الإنترنت، وفي اتساع التغطية على مدى السنوات القليلة القادمة، غير أننا لا يمكن أن نتنبأ كيف يمكن للمكتبات الرقمية أن تستثمر هذا الأداء، لكنه من غير شك سيقدم فرصاً متميزة للغاية.

آفاق البحث والتطوير :

تشكل المكتبات الرقمية الآن مجالاً راسخاً للبحث العلمي، مع ما يتوافر من مستلزمات ورش العمل والمؤتمرات العلمية. وقد كانت هناك محاولات لإصدار دوريات مطبوعة عن المكتبات الرقمية، والأهم من ذلك أن ثمة ما يقرب من ألف شخص على الأقل يعتبرون أن وظائفهم تتمثل في إنجاز البحوث المتصلة بهذا المجال، فعلى سبيل المثال تضمن هذا الكتاب في ثنايا صفحاته نماذج كثيرة لمشروعات تولتها المؤسسة الوطنية للعلوم وهيئة مشروعات البحوث المتقدمة للدفاع المعروفة بداربا DARABA كما نعمت الكثير من المشروعات الحديثة وعلى نحو صريح بالتمويل باعتبارها أبحاثاً خاصة بالمكتبات الرقمية.

ملحوظة ختامية :

لقد اطلعت في رحلة تأليفي لهذا الكتاب على المئات من المصادر التي

كانت الغالبية العظمى منها عبارة عن مواد أولية تمثلت في جهود وصفية كتبها باحثون ومنشؤو مكتبات رقمية لوصف جهودهم في هذا الصدد، وكان أحد هذه المصادر معرض أقيم في رحاب مكتب حق المؤلف في الولايات المتحدة الأمريكية، وآخر عن الكتب التي نفدت من سوق النشر، كما أرسلت الكثير من الرسائل عبر البريد الإلكتروني للكثير من الأصدقاء بشأن بعض الموضوعات، وقد كانت الإنترنت هي مصدري للكثير من الأشياء الأخرى حيث لا تتوافر الكثير من المعلومات عن جوانب ذات علاقة بالموضوع في أشكال تقليدية، وحقاً تعد الإنترنت هي المكتبة الحقيقية لمن يكتب عن المكتبات الرقمية.

إن حلم مكتبات المستقبل يمزج كل شيء نقدره بشأن الطرق التقليدية مع أفضل ما يمكن أن تقدمه المعلومات المباشرة على الإنترنت، وإن كان الخلط بين الجوانب السيئة لكلا الجانبين قد يجد طريقه لبعض الأحلام المزعجة. وإذا كانت الروح الخيرة لأندرو كارنيجي قدمت المكتبات العامة للولايات المتحدة في السنوات الأولى من القرن العشرين، فإن شكلاً جديداً من المكتبات بزغ نجمه الآن. ويحدونا الأمل أن تحظى المكتبات الرقمية بالقدر نفسه من مشاعر الحب والاحترام، وأن تعمل على تلبية المتطلبات العميقة التي صاحبت أحسن أنواع المكتبات ودور النشر منذ زمن طويل.

اللوحة رقم (١٤–١)

ورشة عمل سانتافي **The Santa FEE workshop**

في شهر مارس من عام ١٩٩٧م، قامت المؤسسة الوطنية للعلوم بإقامة

ورشة عمل في سانتافي نيومكسيكو لمناقشة اتجاهات البحوث المستقبلية في مجال المكتبات الرقمية، وكان هذا الأمر جزءاً من عملية تخطيط أدت فيما بعد إلى الإعلان عن برنامج بحثي ضخم جديد خاص بالمكتبات الرقمية، وهو ما تمثل في الفقرة الثانية من مبادرة المكتبات الرقمية. وكانت ورشة العمل هذه فرصة مواتية ليعرض فيها الباحثون تصوراتهم نحو تطوير المكتبات الرقمية ومجالات البحث في هذا المجال.

وقد كان كثير من الحضور في هذا الاجتماع من المرتبطين بمبادرة المكتبات الرقمية أو من المشروعات البحثية الأخرى التي تم تمويلها من قبل الاتحادات الفيدرالية، وبالرغم من أن الحضور كانوا مهتمين بمواصلة بحوثهم، فإنهم لم يوصوا باستمرارية برامجهم البحثية نفسها. لقد كونت المشروعات الأولى مجموعات اختبارية، ثم استخدمتها لأغراض البحث في الموضوعات الفنية بشكل أساسي بعد ذلك. وقد رأى بعض المشاركين في ورشة العمل هذه أن الاستثمار القيم للاعتمادات المالية يجب أن يخصص لبناء مكتبات رقمية كبرى، هذا في الوقت الذي اتفق فيه آخرون أن الحفظ الأرشيفي أمر يستحق قدراً كبيراً من البحث الجاد، ومع ذلك فإن جُل المناقشات انصبت على العمل على جعل المجموعات القائمة أكثر قابلية للاستخدام. ومن الملاحظ أن المشاركين في ورشة العمل كانوا من كبار الباحثين الذين ليسوا في حاجة للعثور على المعلومات فحسب، بل هم في حاجة للتحرر من المعلومات الزائدة عن الحد، وقد أشير في سياق المناقشات أن المشكلة الأساسية لأبحاث المكتبات الرقمية كانت تمثل في فرط المعلومات أو كثرتها فكيف يمكن استخدام الوسائل الآلية لغربلة المعلومات

واستخلاصها ودمجها أو تعزيز مكانتها. وقد تبنت المناقشات طرقاً يستطيع الأفراد من خلالها أن يتمكنوا من إدارة مكتباتهم الخاصة بشكل ناجح، وأن تتمكن المجموعات من القيام بأعمال تعاونية على الإنترنت، فكيف إذن يمكن إدارة المكتبات الرقمية - التي تعني بإدارة مجموعات من المعلومات - لتصبح مرضية للمستفيدين ؟

وقد نُوقشت أيضاً قضايا اجتماعية واقتصادية وقانونية، وكما هو الحال في مثل هذه المجالات، فقد أثير التساؤل حول كيفية التعبير عن استراتيجية بحث متماسكة. وفي الوقت الذي لا يمكن لأحد أن ينكر فيه أهمية هذه المجالات، يظل هناك تشكك بشأن إمكانية معالجة مثل هذه المجالات من قبل مشروعات البحوث الضخمة.

ومن جهة أخرى أشار الفنيون من الحضور في هذا اللقاء إلى أن المكتبات الرقمية باتت من أكبر المؤسسات التي تستخدم الحاسبات العملاقة، وفي عام ١٩٨٦م حينما تأسست مراكز الحاسبات العملاقة الوطنية، كان الكابل الرئيسي للإنترنت ينقل المعلومات بسرعة ٥٦ كيلو بت في الثانية، وكان يشترك فيه كل المستفيدين، ومن الملاحظ أن هذه السرعة توفرها الآن أجهزة المودم رخيصة الثمن التي يستخدمها الأفراد للاتصال بالإنترنت من خلال الاتصال الهاتفي، كما تستطيع الحاسبات المحمولة الآن القيام بما كان لا يقدر على أدائه سوى الحاسبات العملاقة قبل اثني عشر عاماً مضت، ومن المتوقع أن تقوم الحاسبات الأصغر من تلك المحمولة بعد اثني عشر عاماً من الآن بأداء ما تقوم به الحاسبات العملاقة الآن.

وفي غضون سنوات قليلة لاحقة، سيضيف النمو التدريجي السريع إلى تلك التغيرات الأساسية، وقد طلب من المشاركين في ورشة عمل (سانتافي) استشراف الفرضيات الخاصة بالمكتبات الرقمية التي تأصلت بشكل كبير لدرجة أنها أصبحت من المسلمات، وكان الهدف من هذا التحدي هو وضع جدول أعمال خلاق للجيل القادم من بحوث المكتبات الرقمية.

معجم المصطلحات

اشتق مجال المكتبات الرقمية مصطلحاته من مجالات أخرى كثيرة مثل: الحاسب الآلي، والمكتبات، والنشر، والقانون. وسيقدم هذا المعجم توضيحًا مبسطًا للمصطلحات المتنوعة كما هي مستخدمة في بيئة المكتبات الرقمية في وقتنا الحالي. وقد تحمل بعض هذه المصطلحات معاني مختلفة في المجالات الأخرى.

القواعد الأنجلو الأمريكية للفهرسة AACR2 (Anglo-American Catalogin Rules) :

هي مجموعة من القواعد التي تصف محتوى تسجيلات فهارس المكتبات .

خدمات الاستخلاص والتكشيف Abstracting and Indexing Services :

هي خدمات معلومات ثانوية توفر إمكانية البحث في المعلومات البحثية والعلمية، وخاصة عن مقالات الدوريات.

إدارة الوصول للمعلومات Access Management :

هي ضبط عملية الوصول لمواد المكتبات الرقمية، وأحياناً يطلق عليها ضوابط وشروط الوصول للمعلومات، أو إدارة الحقوق .

المكتبة الرقمية لجمعية مصنعي الحاسبات ACM Digital Library :

هي مكتبة رقمية للدوريات ووقائع بحوث المؤتمرات التي نشرتها جمعية مصنعي الحاسبات.

مكتبة الإسكندرية الرقمية Alexandria Digital Library :

هي مكتبة رقمية للمعلومات المكانية ،ومقرها جامعة كاليفورنيا في سانتا باربرا.

الذاكرة الأمريكية وبرنامج المكتبة الرقمية الوطنية

American Memory and the National Digital Library Program :

هي المكتبة الرقمية لمواد مكتبة الكونجرس التي حولت من موادها المصدرية الأولية ذات الصلة بالتاريخ الأمريكي.

برامج الأبليت Applet :

برامج آلية صغيرة يمكن نقلها من الحاسبات الخادمة إلى الحاسبات العميلة، ومن ثم تشغيلها على تلك الحاسبات العميلة.

الأرشيفات Archives :

هي المجموعات وما يصاحبها من أنظمة وخدمات، والتي يتم تنظيمها لأغراض الحفظ طويل الأمد للمعلومات.

مكنز الفنون والعمارة Art and Architecture Thesaurus :

هو مكنز متخصص في مجالات: الفنون الجميلة، والعمارة، وفنون الديكور، والثقافة المادية، وقد طور هذا المكنز في سياق مشروع لشركة جي باول جيتي تراست J. Paul Getty Trust.

منتج مادي Artifact :

هو كيان مادي يتواجد في مكتبة، أو أرشيف، أو متحف.

الترميز المعياري الأمريكي لتبادل المعلومات (آسكي)

ASCII (American Standard Code for Information Interchange) :

هي خطة ترميز يتم بموجبها تمثيل كل حرف بسبعة بتات أو ثمانية؛ وتعد رموز آسكي القابلة للطباعة مجموعة فرعية من الترميز المعياري الأمريكي لتبادل المعلومات.

التحقق من الهوية أو الشخصية Authentication :

المصادقة على أحد المستفيدين، والحاسبات الآلية، أو بعض الكيانات الرقمية، والتأكد من هويته.

الترخيص بالاستخدام Authorization :

منح المستخدم أو الحاسب العميل إذن أو حق الوصول لمعلومات محددة، وتنفيذ أعمال متفق عليها.

تكشيف آلي Automatic Indexing :

إنشاء تسجيلات فهرس أو كشاف عن طريق برامج آلية ، وليس عن طريق المفهرسين .

بحث بوليني أو منطقي Boolean Searching :

أساليب استرجاع المعلومات عندما يتكون التساؤل أو استراتيجية البحث من مصطلحات بحثية، يتم ربطها بروابط، مثل: and، or، not".

متصفح أو مستعرض Browser :

واجهة تفاعل عامة الأغراض تستخدم مع الويب، وغيرها من خدمات المعلومات الأخرى المتاحة على الخط المباشر، وتعرف كذلك بمتصفحات الويب.

تصفح Browse :

اسكتشاف المعلومات اعتمادًا على تنظيم المجموعات، أو من خلال مسح القوائم، هذا فضلاً عن البحث المباشر.

ذاكرة خفية أو خبيئة Cache :

مستودع مؤقت يستخدم للاحتفاظ بنسخة جاهزة من البيانات التي تم استخدامها منذ فترة وجيزة، أو الاحتفاظ بأية بيانات يتوقع أن يتواتر استخدامها بشكل متكرر.

مكتبة كاليفورنيا الرقمية California Digital Library :

مكتبة رقمية تقدم خدماتها لتسعة تجمعات تابعة لجامعة كاليفورنيا.

فهرس Catalog :

مجموعة من التسجيلات الببليوجرافية التي تم إنشاؤها وفقاً لمجموعة مقننة من القواعد.

التصنيف Classification :

تنظيم مواد المكتبة بشكل هرمي حسب فئاتها الموضوعية.

الحاسب العميل Client :

حاسب آلي يعمل نيابة عن المستفيد أو لصالحه، ويشمل ذلك الحاسبات الشخصية للمستفيدين، أو أي حاسب آخر يظهر للحاسب الخادم، ليقوم بتلك المهمة.

واجهة بوابة المرور العامة أو المشتركة CGI (Common Gateway Interface) :

واجهة برمجة تمكن متصفحات الويب من العمل كواجهة

لخدمـات المعلومـات، فضلاً عن مواقع الويب.

ائتلاف المعلومات الشبكية (CNI) Coalition for Networked Information :

ائـتلاف شـريك لجمعيـة المكتبات البحثية والتعليمية، يتعاونـان في مجال المعلومات الشبكات الأكاديمية.

كيان مركب Complex Object :

كيان مكتبـي يتكون مـن مجمـوعة عناصـر ذات علاقات متداخلة، أو من كيانات رقمية.

الضغط Compression :

تقليص لحجم المـواد الـرقمية عن طـريق حـذف المعلـومات الهامشية أو المكـررة redundant، أو عـن طـريق التقـريـب approximation؛ ويمكـن أن يتم عكس الـضغط المحكم lossless - أي إعادة بينية البيانات الأصلية كما كانت قبل الـضغط، في حين لا يمكن الاستعادة في حالة الـضغط غير المحكم lossy بسبب فقدان المعلومات أثناء عملية التقريب.

اللغويات الحاسوبية Computational linguistics :

أحد فروع معالجة اللغة الطبيعية، والتي تتعامل مع قواعد اللغة ولـغوياتـها.

اللغات (المصطلحات) المقيدة Controlled vocabulary :

مجمـوعة مـن المـصطلحات الموضـوعية، مـع قواعد استخدامها عند تعيين المصطلحات المناسبة للمحتوى الموضوعي لأغراض التكشيف والاسترجاع.

تحويل Conversion :

نقل المعلومات المحملة على وسيط إلى وسيط آخر أو تحويلها، بما في

ذلك نقلها من الوسائط الورقية إلى الوسائط الرقمية.

كوربا CORBA :

معيار للعمليات الحاسوبية الموزعة أو اللامركزية حيث يقوم أحد الكيانات المتواجد في أحد الحاسبات بمناشدة وسيط طلب الكيان object Request Broker (ORB) بالتفاعل مع أحد الكيانات الأخرى المتواجدة في حاسب آلي آخر.

مشروع كور CORE :

مشروع لتحويل دوريات متخصصة في الكيمياء من الشكل التقليدي إلى الشكل الرقمي، نفذ خلال الفترة من عام ١٩٩١- ١٩٩٥م من قبل كل من بيل كور Belcore، وجامعة كورنيل، و أو سي إل سي OCLC، والجمعية الأمريكية الكيميائية.

شيفرة Cryptolope :

حاوية آمنة طورتها شركة آي بي إم IBM، تستخدم لشراء محتويات المواد وبيعها على الإنترنت بطريقة آمنة.

نماذج عرض الصفحة المنسدلة CSS (Cascading Style Sheets) :

نظام لعرض الصفحات يستخدم مع لغة ترميز النص الفائق HTML، وهو يعتبر أساس لغة الكتابة الموسعة XLS.

مشروع التقارير الفنية لعلوم الحاسب

CSTR (Computer Science Technical Reports Project) :

مشروع قامت بتمويله هيئة الدفاع للمشروعات البحثية المتقدمة (داربا) ، بالتعاون مع مؤسسة مبادرات البحث الوطني Corporation for National

Research Initiatives، بالإضـافة إلـى خمـس جامعـات هـي: جامعـة كارنيجـي ميلون، وكورنيل، ومعهد إم آي تـي MIT، وستانفورد كاليفورنيا فـي بيركلـي، خلال الفترة من ١٩٩٢ إلى ١٩٩٦م.

هيئة الدفاع للمشروعات البحثية المتقدمة

DARPA - Defense Advanced Research Projects Agency :

أكبر هيئة دولية داعمة لنشاط البحث في مجال علوم الحاسب بما في ذلك مجال المكتبات الرقمية في الولايات المتحدة الأمريكية، وكانت تعرف سابقاً بأربا ARPA.

نوع البيانات Data type :

ما وراء بيانات بنائية ترافق البيانات الرقمية لتشير إلى الصيغة الرقمية أو إلى التطبيق المستخدم لمعالجة البيانات.

معيار تشفير البيانات DES (Data Encryption Standard) :

وسيلة للتشفير بالمفتاح الخاص private-key encryption.

تصنيف ديوي العشري Dewey Decimal Classification :

خطة لتصنيف مواد المكتبة تستخدم الترميز الرقمي للتعبير عن المجالات الموضوعية.

تصور سطح المكتب Desktop metaphor :

تصور واجهة المستفيد على الحاسبات الشخصية، يمثل المعلومات كملفات ومجلدات على سطح المكتب.

داينست Dienst :

بنية فنية لخدمات مكتبة رقمية، وبروتوكول مفتوح يوفر تلك الخدمات، وقد طور هذا النموذج في جامعة كورنيل، واستخدم من قبل المكتبة المرجعية الفنية الشبكية لعلوم الحاسب NCSTRL.

علم الآثار الرقمي Digital Archeology :

عملية استرجاع المعلومات من مصادر البيانات التالفة، وغير المكتملة، والقديمة.

مبادرة المكتبات الرقمية Digital Libraries Initiative :

برنامج بحثي خاص بالمكتبات الرقمية، وخلال المرحلة الأولى من هذا البرنامج والتي امتدت خلال الفترة من ١٩٩٤- ١٩٩٨م، مولت كل من المؤسسة الوطنية للعلوم، وداربا، وناسا، ستة مشروعات جامعية، وقد بدأت المرحلة الثانية في هذا البرنامج في عام ١٩٩٨م.

كيان رقمي Digital Object :

عنصر يختزن في مكتبة رقمية، ويتكون من البيانات، وما وراء البيانات، والمحدد.

توقيع رقمي Digital Signature :

رمز مشفر مكون من قيمة اختبارية hash يستخدم للإشارة إلى أن البيانات لم يتم تغييرها، وذلك الرمز يشفر مع المفتاح العام لمنشئ التوقيع .

البث Dissemination :

عملية النقل للشكل المختزن به الكيان الرقمي في الحاسب المستودع إلى الحاسب العميل.

العمليات الحاسوبية الموزعة أو اللامركزية Distributed Computing :

نظم حاسوبية تقدم الخدمات عن طريقها للمستفيدين من خلال مجموعة من الحاسبات المترابطة والموصلة على الشبكة .

صحيفة المكتبات الرقمية D-Lib Magazine :

صحيفة شهرية متاحة على الخط المباشر على الإنترنت تهتم بالأبحاث في مجال المكتبات الرقمية والابتكارات الجديدة فيه.

دي ليت DLITE :

واجهة تفاعل تجريبية استخدمت مع نظام جامعة ستانفورد المسمى باستانفورد إنفو بص Stanford University InfoBus

وثيقة Document :

كيان رقمي مناظر للوثيقة المادية، وخاصة المواد النصية؛ ويعتبر نموذج الوثائق أو نموذج الوثيقة document model نموذجًا كيانيًا للمستندات.

اسم النطاق Domain Name :

هو اسم الحاسب على الإنترنت، ويقوم نظام خدمة اسم النطاق (DNS) بتحويل أسماء النطاق إلى عناوين بروتوكول الإنترنت IP.

محدد الكيان الرقمي أو معرفه DOI (Digital Object Identifier) :

محدد يستخدمه الناشرون لتحديد هوية المواد التي تنشر إلكترونياً، وهو شكل من أشكال المعالجات (الهاندلز handle).

لغة العناصر الدلالية لأسلوب الوثيقة ومواصفاتها

DSSSL (Document Style Semantics and Specification Language) :

نظام ذو هدف عام لنماذج عرض الصفحات عند ترميزها باستخدام اللغة المعيارية الموحدة لترميز النصوص SGML.

معرف نوع الوثيقة DTD (Document Type Definition) :

مواصفة ترميز أو تهيئة لمجموعة أو فئة من المستندات المعرفة داخل إطار اللغة المعيارية الموحدة لترميز النصوص SGML.

دبلن كور Dublin Core :

مجموعة مبسطة من عناصر ما وراء البيانات المستخدمة في المكتبات الرقمية، وتتمثل وظيفتها الرئيسية في وصف الكيانات الرقمية، ولأغراض إدارة المجموعات، وتبادل ما وراء البيانات.

كيان نشط أو حركي Dynamic Object :

كيان رقمي يتواجد عندما يتم عرض البث للمستفيد اعتماداً على تنفيذ برنامج آلي، أو نشاط خارجي آخر .

الوصف الأرشيفي المرمز EAD (Encoded Archival Description) :

أحد معرفات أنواع الوثائق يستخدم لترميز الإصدارات الإلكترونية من معينات البحث للمواد الأرشيفية.

دورية إلكترونية Electronic Journal :

مطبوع متاح مباشرة ومنظم كالدورية التقليدية المطبوعة، وقد تكون هذه الدورية الإلكترونية عبارة عن إصدارة أخرى من دورية مطبوعة، أو دورية إلكترونية الأصل أي ليس لها مناظر مطبوع.

المكتبة الإلكترونية (إي لب) eLib :

برنامج بريطاني للابتكارات حول موضوع النشر الإلكتروني .

المحاكاة Emulation :

إعادة نسخ نظام تحسيب آلي أو تكراره، بهدف معالجة برامج وبيانات مأخوذة من نظم قديمة لم تعد متوافرة الآن.

تشفير Encryption :

مجموعة أساليب فنية تتبع لترميز المعلومات لإحاطتها بقدر من الخصوصية والأمان، لذا تبدو تلك المعلومات في صورة بيانات مبعثرة. وتتطلب عملية إعادة البيانات لحالتها الأصلية - المعروفة بفك التشفير - معرفة المفتاح الرقمي.

كيانات، عناصر Entities, Elements :

في لغة ترميز النصوص، تعد الكيانات entities هي الوحدة الأساسية للمعلومات، بما في ذلك الكيانات الحرفية؛ في حين تكون العناصر elements عبارة عن سلسلة من الكيانات التي تشكل وحدة بنائية.

تعبير Expression :

تجسيد حقيقة العمل من خلال التعبير عن المفهوم المجرد في شكل كلمات فعلية، أو في شكل أصوات، أو صورة أو غير ذلك من أشكال التعبير .

استخدام مشروع Fair Use :

أحد مفاهيم قانون حقوق التأليف، والذي يسمح باستخدام نسخ محددة من

المواد المنشورة من غير الحصول على تصريح بذلك من المالك أو صاحب حق التأليف، كأن يكون لذلك لأغراض البحث العلمي أو المراجعة العلمية.

المكتبة الرقمية الاتحادية Federated Digital Library :

مجموعة من المكتبات الرقمية تدعم المعايير والخدمات المشتركة، ومن ثم توفر التشغيل المتداخل بين هذه المكتبات، وتقدم الخدمة المترابطة للمستفيدين.

حقول، حقول فرعية Field, Subfield :

عناصر مستقلة للمعلومات موجودة في تسجيلات بنائية، مثل: تسجيلة الفهرس، أو تسجيلة قاعدة البيانات.

بحث حقلي Fielded Searching :

أساليب بحث المواد النصية، بما في ذلك الفهارس؛ حيث تتم مضاهاة المصطلحات البحثية بمحتويات حقول محددة.

معينات البحث Finding Aid :

وثيقة نصية تصف مقتنيات أحد الأرشيفات، أو المكتبات، أو المتاحف.

الجدار الناري Firewall :

نظام حاسوبي يقوم بمراقبة البيانات الممررة عبر أجزاء الشبكة، ويستخدم لتوفير نظام أمن للشبكة المحلية الخاصة عند نقطة الاتصال بالإنترنت.

البيع أولاً First Sale :

أحد مفاهيم قانون حقوق التأليف، والذي يعطي لمشتري الكتاب أو أي

كيان آخر حق نقله لشخص آخر، دون الحاجة إلى تصريح أو إذن من صاحب حقوق التأليف.

بروتوكول نقل الملفات FTP (File Transfer Protocol) :

بروتوكول يستخدم لنقل الملفات بين الحاسبات على شبكة الإنترنت.

بحث النصوص الكاملة Full text searching :

أساليب بحث المواد النصية؛ حيث يتم مضاهاة الاستفسار أو استراتيجية البحث بالنص الكامل للمادة العلمية.

الجامع Gatherer :

برنامج يقوم آلياً بتجميع المعلومات التكشفية من مجموعات المكتبة الرقمية.

معجم Gazetteer :

هي قاعدة بيانات تستخدم للترجمة والتنقل بين العروض المختلفة للمراجع المكانية: مثل أسماء الأماكن، والإحداثيات الجغرافية.

الأنواع أو الأجناس الأدبية Genre :

مجموعة أو فئة لأحد الكيانات عندما ينظر لهذا على أنه عمل فكري.

المعلومات المكانية Geospatial Information :

هي المعلومات المرجعية للأماكن الجغرافية.

جف GIF :

صيغة لحفظ الصور المضغوطة.

جوجل Google :

برنامج لبحث الويب يقوم بترتيب صفحاتها ترتيباً طبقياً في شكل قائمة، وذلك من خلال إعطاء وزن للارتباطات التي تشير إلى صفحة محددة.

جوفر Gopher :

بروتوكول تم تطويره قبل الويب لبناء المكتبات الرقمية، ولم يعد مستخدماً الآن.

المعالجة (هاندلز) Handle :

نظام تسمية فريد وعالمي لمصادر الإنترنت وأنظمة الحاسب بهدف إدارتها، وقد طورت هذا النظام ائتلاف مبادرة البحث الوطني Coalition National Research Initiative ، كشكل من أشكال نظام الأسماء الموحدة للمصادر URN.

هارفست Harvest :

مشروع بحثي انتهى إلى تطوير بنية فنية للبحث الموزع أو اللامركزي، بما في ذلك البروتوكولات والصيغ.

القيمة الاختبارية (هاش) Hash :

قيمة قصيرة يتم حسابها أو إنشاؤها من البيانات الرقمية، وتهدف إلى التمييز بين تلك البيانات وأية بيانات أخرى [لأغراض اكتشاف تزوير البيانات].

مطبعة هاي واير HighWire Press :

مجازفة تجارية في مجال النشر قامت بها مكتبات جامعة ستانفورد، وتمثلت في إتاحة نسخ إلكترونية من الدوريات لحساب المجتمعات التعليمية والمهنية.

إصابة Hit :

١- الطلب أو التساؤل المقدم لخادم الويب أو أي نظام آلي آخر .

٢ - في سياق استرجاع المعلومات: هي الوثيقة التي تم استكشافها، ومن ثم استرجاعها استجابة لتساؤل المستفيد.

صفحة رئيسية Home Page :

الصفحة الأولى أو التمهيدية لمجموعة من المعلومات المتاحة على الويب .

لغة تهيئة النصوص الفائقة أو ترميزها

HTML (Hypertext Markup Language) :

لغة مبسطة لترميز المواد النصية وتنسيقها أو تشكيلها، وربط هذه النصوص بالكيانات الأخرى، وهي تستخدم مع الويب.

بروتوكول نقل الملفات الفائقة HTTP (Hypertext Transfer Protocol) :

البروتوكول الأساسي للويب والمستخدم في عملية الاتصال بين متصفحات الويب ومواقع الإنترنت.

رابط فائق Hyperlink :

رابط شبكي من عنصر موجود في مكتبة رقمية أو متاح على أحد مواقع الويب إلى عناصر أخرى.

الاتحاد الدولي للبحث في مجال السياسة والعلوم الاجتماعية

ICPSR (International Consortium for political and Social Science Research) :

أرشيف مقره جامعة ميتشجان لمجموعات البيانات في مجال العلوم الاجتماعية.

محدد أو معرف Identifier :

سلسلة من الحروف تحدد هوية مصدر معين موجود في مكتبة رقمية أو متاح على الشبكة.

لجنة العمل الخاصة بهندسة الإنترنت

IETF (Internet Engineering Task Force) :

الجهاز الذي يضطلع بمهام التنسيق بين جوانب التطوير الفنية للإنترنت بما في ذلك المقاييس أو المعايير.

إنفوباص InfoBus :

أحد مشروعات التشغيل المتداخل التي تستخدم الملقمات كواجهات تعامل بين النظم العاملة، وقد طورت هذا المشروع جامعة ستانفورد.

استكشاف المعلومات Information Discovery :

مصطلح عام يشمل جميع استراتيجيات وأساليب إيجاد المعلومات في المكتبة الرقمية.

استرجاع المعلومات Information Retrieval :

بحث مجموعة من المعلومات عن كيانات تضاهي تساؤلاً أو استراتيجية بحث.

إنفورميديا Informedia :

برنامج بحثي ومكتبة رقمية لمواد الفيديو اتخذ من جامعة كارنيجي ميلون مقراً لـه.

إنسبك Inspec :

خدمة تكشيف متخصصة في مجالات: الفيزياء، والهندسة، وعلوم الحاسب الآلي، والمجالات الأخرى ذات الصلة.

الإنترنت Internet :

شبكة عالمية تتألف من مجموعة من الشبكات التي تدار كل منها بشكل مستقل، وتستخدم بروتوكول التحكم بالإرسال / بروتوكول الإنترنت (TCP/IP) ، ونظام تسمية مشترك. وهي بمثابة وريث لشبكة آربا ARPAnet.

سلسلة طلب التعليق الخاصة بالإنترنت Internet RFC Series :

هي المطبوعات الفنية للإنترنت والتي تصدرها لجنة العمل الخاصة بهندسة الإنترنت. وتعتبر مسودة الإنترنت نسخة أولية من طلبات التعليق.

تشغيل متداخل (بيني) Interoperability :

مهمة بناء خدمات مترابطة للمستفيدين من مجموعة مكونات، مختلفة من حيث بنيتها الفنية، ومستقلة من حيث بنيتها الإدارية.

ملف مصنف Inverted File :

قائمة بالكلمات الموجودة في مجموعة من الوثائق مع تحديد أماكن

وجودها في تلك الوثائق، أما القائمة المصنفة فهي القائمة التي تحوي مكان وجود كلمة معينة.

عنصر أو مفردة بيانات Item :

جزء محدد من مواد المكتبة الرقمية، وهو يعتبر حالة أو نسخة فردية من حالات أو عروض الحالات التي تكون عليه البيانات.

جافا Java :

لغة برمجة تستخدم في كتابة الرمز المتنقل، وخاصة لأغراض واجهات المستفيدين. وقد طورت هذه اللغة شركة صن مايكروسيستمز Sun Microsystems.

جافا سكريبت JavaScript :

لغة نصية أو لغة نصوص تستخدم لتضمين التعليمات القابلة للتنفيذ على إحدى صفحات الويب.

صيغة جي بيج JPEG :

صيغة لحفظ الصور المضغوطة.

جستور JSTOR :

هو مشروع يقدم خدمة الاشتراكات تبنته مؤسسة أندرو ميلون بهدف تزويد المكتبات الأكاديمية بالأعداد القديمة من الدوريات الهامة.

مفتاح Key :

رمز رقمي يستخدم لتشفير الرسائل وإعادة فك التشفير. ويستخدم تشفير المفتاح الخاص مفتاحًا سريًا واحدًا، في حين يستخدم تشفير المفتاح المزدوج

(أو المفتاح العام) مفتاحين أحدهما سري والآخر عام.

نظام موروث Legacy System :

نظام موجود، عادة ما يكون نظاماً آلياً، وهو الذي يجب أن يكن مجهزاً أو مكيفاً لبناء نظم جديدة .

معجم Lexicon :

أداة لغوية تتضمن معلومات عن الصيغ المورفولوجية المتغيرة، والاستخدامات النحوية للكلمات.

لكزس Lexis :

خدمة معلومات قانونية، وهي الخدمة الرائدة لإتاحة النصوص الكاملة على الخط المباشر في مجال القانون .

أرشيفات لوس ألاموس للطباعة الإلكترونية Los Alamos E-Print Archives :

موقع متاح على الخط المباشر لأغراض النشر السريع للأوراق البحثية في علم الفيزياء، والتخصصات العلمية الأخرى ذات الصلة.

عرض البيانات أو عرض الحالة Manifestation :

شكل يعطى للتعبير عن حالة العمل، كأن يتم عرض العمل في شكل رقمي على سبيل المثال.

مارك (الفهرسة المقروءة آلياً) MARC (Machine-Readable Cataloging) :

صيغة تستخدمها المكتبات لاختزان تسجيلات الفهرسة وتبادلها.

لغة تهيئة أو ترميز Markup Language :

رموز متضمنة في وثيقة ما لتصف بنية هذه الوثيقة والشكل الذي ينبغي أن تظهر فيه، أو لتصف كليهما.

مدلاين Medline :

خدمة تكشيف للبحوث في مجال الطب ، والمجالات المرتبطة به، وتقدم هذه الخدمة المكتبة القومية الطبية [في الولايات المتحدة الأمريكية].

ملفيل MELVYL :

مكتبة رقمية مشتركة للهيئات الأكاديمية في ولاية كاليفورنيا؛ وهي جزء من مكتبة كاليفورنيا الرقمية.

ممكس Memex :

تصور لمكتبة مباشرة اقترحه فانيفار بوش في عام ١٩٤٥م.

مركوري Mercury :

مشروع مكتبة رقمية تجريبية لإتاحة الدوريات العلمية على الخط المباشر في جامعة كارنيجي ميلون، وتم ذلك خلال الفترة من ١٩٨٧-١٩٩٣ م.

مش (قائمة رؤوس الموضوعات الطبية) Me SH (Medical Subject Headings) :

مجموعة من المصطلحات الموضوعية يرافقها مكنز، تستخدم لوصف البحوث الطبية، وتبنت المكتبة القومية الطبية إعداد هذه القائمة وصيانتها.

ما وراء البيانات (الميتاداتا) Metadata :

بيانات تصف البيانات، وعادة تنقسم إلى الفئات التالية: ما وراء البيانات الوصفية مثل البيانات الببليوجرافية؛ وما وراء البيانات البنائية، وهي المعلومات عن أشكال البيانات وهياكلها البنائية؛ وما وراء البيانات الإدارية والتي تستخدم لإدارة المعلومات.

نقل Migration :

حفظ المحتويات الرقمية، حيث يحتفظ بالمعلومات الأساسية، ويتم إعادة نسخ الصيغ القديمة والهياكل الداخلية في صيغ جديدة .

مايم (نوع وسيط الإنترنت) MIME (internet media type) :

خطة لتوصيف نوع بيانات المادة الرقمية.

تكرار أو مطابقة Mirror :

نظام حاسوبي يحتوي على نسخة مكررة أو مطابقة من المعلومات المختزنة في نظام آخر.

رمز متنقل Mobile Code :

برامج حاسوبية أو أجزاء من برامج حاسوبية يتم نقلها عبر شبكة، ويقوم الحاسب البعيد بتنفيذها .

علم الصرف Morphology :

الأشكال المتغيرة النحوية وغيرها من الأشكال التي تتخذها الكلمة، والتي تشتق من جذر الكلمة نفسه.

موزايك (الفسيفساء) Mosaic :

أول متصفحات الويب التي شاع استخدامها بشكل كبير، وقد قامت

جامعة إلينويز في أربانا بتطوير هذا المتصفح .

إم بيج MPEG :

عائلة من صيغ ضغط واختزان المواد الفيديوية والصوتية المرقمنة .

وسائط متعددة MultiMedia :

مزيج من عدة أنواع وسائط مختلفة موجودة في كيان رقمي واحد، أو مجموعة كيانات، مثال ذلك: الصور، والصوت، والفيديو.

معالجة اللغة الطبيعية Natural Language processing :

استثمار إمكانات الحاسبات في تفسير الكلمات أو ترجمتها ومعالجتها على اعتبار أنها جزء من مكونات اللغة.

المكتبة المرجعية الفنية الشبكية لعلوم الحاسب

NCSTRL (Networked Computer Science Technical Reference Library)

هي مكتبة عالمية لا مركزية للمواد المتخصصة في علوم الحاسب وخدماتها، وقد اتخذت هذه المكتبة من جامعة كورنيل مقرًا لها .

نت لب Netlib :

مكتبة رقمية متخصصة في برمجيات علم الرياضيات والمجموعات المرتبطة به.

المؤسسة القومية للعلوم NSF (National Science Foundation) :

هيئة حكومية أمريكية تقوم بدعم الأبحاث في مجال العلوم والهندسة، بما

في ذلك البحوث المتخصصة في المكتبات الرقمية.

كيان Object :

مصطلح حاسوبي فني لجزء مستقل من رمز الحاسب، مع البيانات المرتبطة به. ومن ثم جاءت ما تعرف بالبرمجة الموجهة للكيانات، والكيانات الموزعة، حيث تكون الكيانات مترابطة بعضها مع بعض عبر الشبكة.

نموذج كياني Object Model :

وصف العلاقات البنائية أو الهيكلية بين مكونات الكيانات المكتبية، بما في ذلك ما وراء البيانات المرتبطة بها.

أو سي إل سي (مركز الحاسب الآلي للمكتبات على الخط المباشر) :

OCLC (Online Computer Library System) :

منظمة توفر خدمة ببليوجرافية للمكتبات لتقاسم تسجيلات الفهرسة، وذلك من بين خدمات أخرى تقدمها.

أوباك (الفهرس المتاح للجمهور على الخط المباشر) :

OPAC (Online Public Access Catalog) :

فهرس للمكتبة متاح على الخط المباشر يستخدمه جمهور المكتبة.

وصول مجاني أو حر Open Access :

المصادر المتاحة على الخط المباشر للمستفيدين دون الحاجة للحصول على ضمانات التحقق من الشخصية أو دفع مقابل مادي.

التعرف إلى الأحرف بصرياً Optical Character Recognition :

تحويل آلي للنصوص من صورة مرقمنة إلى نص آلي يتعامل معها الحاسب.

باد بلص بلص ++Pad :

واجهة تعامل تجريبية للوصول لمجموعة كبيرة من المعلومات اعتماداً على نظام التكبير والتصغير الدلالي .

لغة وصف الصفحات Page Description Language :

نظام لترميز الوثائق والتي تصف بدقةٍ شكل ظهور الوثيقة عند عملية الطباعة على الورق أو عرضها على الشاشة.

بي دي إف (صيغة الوثيقة القابلة للنقل) PDF (Portable Document Format) :

لغة وصف الصفحات، طورتها شركة أدوبي لاختزان صور الصفحات واسترجاعها.

التحكيم العلمي Peer review :

الإجراء الذي يتم بموجبه عرض مقالات الدوريات العلمية على باحثين آخرين غير مؤلفيها للحكم عليها قبل نشرها .

برسيوس Perseus :

مكتبة رقمية للمصادر المترابطة فيما بينها بروابط فائقة، والمتخصصة في مجال الدراسات الكلاسيكية، والمجالات ذات الصلة. وتأسست هذه المكتبة في جامعة تافتس.

سياسة Policy :

قاعدة أو قانون يضعه مدير مكتبة رقمية يبين فيه هوية المستفيدين المرخص لهم بالوصول لمواد المكتبة، وطبيعة المواد التي يستطيعون الوصول إليها.

منفذ Port :

وسيلة يستخدمها بروتوكول التحكم بالإرسال TCP ليحدد بها أي البرامج المشغلة على الحاسب، والتي يجب عليها معالجة الرسائل القادمة من شبكة الإنترنت.

بوست سكريبت PostScript :

لغة برمجة تستخدم لإنتاج المخرجات الرسومية أو البيانية لأغراض الطباعة، وهي تستخدم كلغة وصف الصفحات.

التحقيق أو الدقة Precision :

تستخدم في سياق استرجاع المعلومات للدلالة على النسبة المئوية للإصابات hits التي تلبي تساؤل المستفيد والذي عبر عنه في استراتيجية البحث.

ملف العرض Presentation Profile :

توجيهات مرافقة للكيان الرقمي تشير إلى الشكل الذي يجب أن يعرض به للمستفيد.

بروتوكول Protocol :

مجموعة قواعد تصف تتابع الرسائل المرسلة عبر الشبكة، محددة بناء جمل هذه الرسائل أو ترتيب كلماتها ودلالاتها.

ملقم Proxy :

حاسب آلي يعمل كحلقة وصل بين نظامين آليين يستخدمان معايير أو صيغًا أو بروتوكولات مختلفة.

النشر Publish :

أتح المعلومات ووزعها على العامة.

محدد موحد ثابت لموقع المصدر PURL (Persistent URL) :

أسلوب لتوفير محددات ثابتة باستخدام بروتوكولات الويب القياسية، وقد طور هذا الأسلوب مركز الحاسب الآلي للمكتبات على الخط المباشر OCLC.

استراتيجية بحث Query :

سلسلة من الكلمات النصية، التي بنيت بشكل احتمالي، لتستخدم في عملية استرجاع المعلومات بهدف إيجاد الكيانات التي تضاهي الكلمات التي تتواجد في الاستراتيجية .

بحث طبقي Ranked Searching :

أسلوب لاسترجاع المعلومات يقوم باسترجاع قائمة من الوثائق مرتبة ترتيباً طبقياً تبعاً لمدى مطابقة كل وثيقة منها لاستراتيجية البحث .

إطار وصف المصدر RDF (Resource Description Framework) :

أسلوب لتحديد بناء ما وراء البيانات، وذلك لأغراض تبادل ما وراء البيانات.

ريال أديو RealAudio :

صيغة وبروتوكول لضغط الملفات الصوتية المرقمنة واختزانها، ونقلها عبر الشبكة ليتم تشغيلها في وقتها الحقيقي.

استدعاء Recall :

في سياق عملية استرجاع المعلومات: هي النسبة المئوية للمواد الصالحة أي التي تلبي حاجة المستفيد من مجموع المواد التي تم بالفعل استرجاعها عند البحث.

تحديث أو تنشيط Refresh :

عمل نسخة طبق الأصل للبيانات من وسيط قديم إلى وسيط جديد لأغراض الحفظ طويل الأجل .

ترجمة Render :

تحويل المعلومات الرقمية من الشكل التي استقبلت به من الحاسب المستودع إلى الشكل الذي تعرض به للمستفيد على شاشة الحاسب، أو أي شكل آخر من أشكال العرض.

إعادة نسخ أو تكرار Replication :

عمل نسخ من المادة الرقمية بهدف النسخ الاحتياطي، أو لتحسين كفاءة الأداء، أو لتقوية مدى موثوقيتها، أو لأغراض صيانتها والحفاظ عليها .

مستودع Repository :

نظام حاسوبي يستخدم لاختزان مجموعات المكتبة الرقمية وبثها للمستفيدين.

التشفير بطريقة ريفست – شامير – أولمان RSA Encryption :

طريقة من طرق التشفير باستخدام المفتاح المزدوج (المفتاح العام).

مسح ضوئي Scanning :

إحدى طرق التحويل التي يتم بموجبها تمثيل الكيانات المادية مثل الصفحة المطبوعة، بشبكة رقمية مربعة من النقاط أو البكسيل Digital grid of pixels.

مصطلح بحث Search Term :

مصطلح فردي يرد ضمن استراتيجية البحث، عادة يكون في شكل كلمة واحدة، أو عبارة قصيرة.

معلومات ثانوية Secondary Information :

مصادر المعلومات التي تصف المعلومات الأخرى (الأولية) مثل: الفهارس والكشافات والمستخلصات المستخدمة في الحصول على المعلومات وإدارة المجموعات.

أمن Security :

الأساليب والممارسات التي تتكفل بحماية صحة نظم الحاسبات وسلامتها، وخدمات المكتبة الرقمية ومجموعاتها.

حاسب خادم Server :

أي حاسب موجود على الشبكة، دون الحاسب العميل (حاسبات المستفيدين) ، والذي يقوم باختزان المجموعات، أو يقدم الخدمات للمستفيدين.

إس جي إم إل (اللغة المعيارية الموحدة لتهيئة أو ترميز النصوص)

SGML (Standard Generalized Markup Language) :

نظام لإنشاء لغات التهيئة أو الترميز التي تمثل بنية الوثيقة.

إس آي سي آي (محدد الوحدات والمساهمات المسلسلة)

SICI (Serial Item and Contribution Identifier) :

محدد أو معرف لإحدى إصدارات دورية معينة، أو إحدى المقالات المنشورة فيها.

التعرف إلى الاستفسارات المنطوقة أو التعرف الصوتي Speech Recognition :

عملية تحويل آلية للكلمات المنطوقة إلى نص حاسوبي.

ستارتز STARTS :

بروتوكول تجريبي مستخدم في البحث الموزع، يساعد الحاسب العميل في دمج نتائج البحث التي يتم الحصول عليها من متصفحات بحث متعددة.

التجذير Stemming :

هي الحد من إرجاع الصيغ المختلفة للكلمة إلى الجذر العام عند استرجاع المعلومات.

وفي سياق استرجاع المعلومات: الحد من استرجاع جميع الصيغ المورفولوجية للكلمة، وقصر ما يتم استرجاعه على ما لـه علاقة بالجذر الشائع للكلمة.

كلمة توقف أو استبعاد Stop Word :

كلمة شائعة الاستخدام بالدرجة التي تدعو لتجاهلها عند عملية استرجاع المعلومات. ويطلق على مجموعة الكلمات الشائعة قائمة التوقف أو

الاستبعاد.

نوع بنائي Structural Type :

ما وراء البيانات تبين الفئة البنائية للكيان الرقمي.

نموذج الكتابة أو عرض الصفحة Style Sheet :

مجموعة قواعد تصف الكيفية التي تترجم بها علامات الترميز التي تتضمن في الوثيقة، إلى المظهر الذي ينبغي أن تعرض به تلك الوثيقة للمستفيد.

اشتراك Subscription :

في المكتبات الرقمية، يتم دفع رسوم الاشتراك للوصول إلى المجموعات أو خدمات محددة إما بصورة فردية أو من خلال مؤسسة ما. وعادة ما تكون الاشتراكات لفترة محددة (مثل اشتراك لمدة عام).

استخدام لاحق أو غير مباشر Subsequent Use :

استخدام المواد الرقمية أو الإفادة منها بمجرد خروجها عن سيطرة المكتبة الرقمية.

تاج Tag :

سلسلة حروف خاصة متضمنة في نص مرمز أو مهيأ بإحدى لغات الترميز لتوضيح بنية هذا النص أو صيغته.

بروتوكول التحكم بالإرسال / بروتوكول الإنترنت TCP/IP :

البروتوكولان الأساسيان المستخدمان في شبكة الإنترنت. حيث يستخدم بروتوكول الإنترنت عناوين بروتوكولات الإنترنت الرقمية للربط أجزاء

الـشبكة، فــي حـين يقـوم بـروتوكول الـتحكم بالإرسـال بتوفيـر الـتوزيع أو الإرسال الآمن للرسائل بين الحاسبات المتصلة بالشبكة.

تي إي آي (مبادرة ترميز النص) TEI (Text Encoding Initiative) :

مشروع لتمثيل النصوص في صور رقمية، مع التركيز على النصوص التي تتصل باحتياجات الباحثين في مجال الإنسانيات، مع استخدام معرفات أنواع الوثائق DTD بواسطة البرنامج هذا المشروع.

تي إكس TeX :

أحد أساليب ترميز النصوص والذي يصف بدقة مظهر تلك النصوص عند طباعتها ، وخاصة النصوص ذات الصبغة الرياضية. وتعد لغة لا تي إكس LaTex إحدى إصدارات لغة تي إكس TeX.

مكنز Thesaurus :

أداة لغوية تربط الكلمات بعضها ببعض بناء على معاني هذه الكلمات.

مدرسة تايسر الصيفية Ticer Summer School :

برنامج تنظمه جامعة تلبيرج الهولندية يهدف إلى تثقيف ذوي الخبرة من المكتبيين بمجال المكتبات الرقمية.

تبستر Tipster :

برنامج للأبحاث تبنته داربا DARPA بهدف الارتقاء بدقة أساليب معالجة النصوص، بما في ذلك عملية استرجاع المعلومات.

النقل الصوتي للحروف Transliteration :

طريقة منهجية لتحويل الحروف المكتوبة بلغة أو تهجئة معينة إلى لغة أو تهجئة أخرى.

تريك (مؤتمر استرجاع النصوص) TREC (Text Retrieval Conference) :

مؤتمرات سنوية يتم فيها تقويم أساليب معالجة النصوص عن طريق اختبارها مقابل مجموعات ومهام قياسية.

بتر Truncation :

استخدام الحروف القليلة الأولى من الكلمة كمصطلح بحثي عند عملية استرجاع المعلومات.

تيولب Tulip :

تجربة تبناها الناشر إلسفير Elsevier، حيث قام بمسح ضوئي لمجموعة من الدوريات العلمية، ثم قامت مجموعة من الجامعات بتحميلها ونشرها على أجهزة الحاسب المحلية لديهم.

يو دي بي UDP :

أحد بروتوكولات الإنترنت يقوم بنقل حزم البيانات دون إجراء فحص الأخطاء.

يونيكود [نظام الرمز الموحد] Unicode :

كود مكون من ١٦ بتًا ليمثل الحروف المستخدمة في معظم الخطوط المكتوبة على مستوى العالم. ويعد نظام يو تي إف - ٨ UTF8 طريقة ترميز بديل، يقوم بموجبها البايت المكون من رمز أو أكثر من ثمانية بتات، بتمثيل كل حرف من حروف نظام اليونيكود.

فهرس موحد Union Catalog :

فهرس واحد يشتمل على تسجيلات تعرف بمقتنيات عدة مجموعات أو مكتبات.

يو آر إل (المحدد الموحد لموقع المصدر) URL (Uniform Resource Locator) :

إشارة لمصدر موجود على الإنترنت، يصف البورتوكول المستخدم، والحاسب، والملف الموجود على هذا الحاسب، ومعالم لهذا الملف أو متغيراته. والمحدد الموحد المطلق أو الثابت يصف مكان المصدر في شكل اسم نطاق أو عنوان بروتوكول الإنترنت؛ أما المحدد الموحد النسبي فيصف مكان المصدر القريب من مكان الملف الحالي.

يو آر إن (الاسم الموحد للمصدر) URN (Uniform Resource Name) :

اسم موقع مستقل لمصادر الإنترنت.

وايز WAIS :

نسخة أولية من بروتوكول Z39.50 استُخدمت في المكتبات الرقمية قبل الويب، ولم يعد مستخدماً في الوقت الحالي.

إطار وارويك Warwick Framework :

نموذج عام يصف الأجزاء المختلفة للكيانات المركبة،بما في ذلك الفئات المختلفة من ما وراء البيانات (الميتاداتا).

علامة مائية Watermark :

رمز ما يتم تضمينه في المادة الرقمية لإثبات ملكية هذه المادة، وقد يكون هذا الرمز مرئياً للمستفيد أو غير مرئي.

زاحف الويب (ويب كرولر) Web Crawler :

برنامج تكشيف على الويب، يقوم ببناء كشاف من خلال تتبع الروابط بصورة مستمرة من صفحة ويب إلى صفحة أخرى.

مدير موقع الويب Webmaster :

الشخص الذي يقوم بإدارة مواقع الويب .

خدمات بحث الويب Web Search Services :

خدمات تجارية تقدم خدمة بحث الويب، من أمثلتها: ياهو، وألتافيستا، وإكس سايت، وليكوس، وإنفوسيك وغيرها.

موقع الويب Web Site :

مجموعة معلومات على الويب، عادة ما يتم اختزانها على خوادم الويب .

ويستلو Westlaw :

خدمة معلومات قانونية قدمتها شركة ويست للنشر.

نسيج الشبكة العالمية (العنكبوتية) World Wide Web (web) :

شبكة النسيج العالمية (الويب) .

هي مجموعة مترابطة من مصادر المعلومات المتوفرة على الإنترنت، والتقنيات المستخدمة معها بما في ذلك لغة تهيئة النصوص الفائقة أو ترميزها، وبروتوكول نقل الملفات الفائقة، والمحددات الموحدة لمواقع المصدر ، ومعيار مايم.

اتحاد شبكة النسيج العالمية World wide web consortium (W3c) :

اتحاد عالمي تأسس في معهد ماساشوسيتس للتقنية، يضطلع بمهام تنسيق جهود التطورات التقنية المتصلة بالويب.

عمل Work :

تجريد فكري أساسي كامن وراء بعض المواد في المكتبة الرقمية.

لغة حقوق الملكية الرقمية (زيروكس)

Xerox Digital Property Rights Language :

هي نصوص وقواعد تعبر بوضوح عن حقوق استخدام الأعمال الرقمية وضوابطها ورسومها.

لغة زيروكس لحقوق الملكية الرقمية

Xerox Digital Property Rights Language :

كلمات وحروف تشير إلى حقوق استخدام الأعمال الرقمية وشروطه ورسومه.

لغة الكتابة الموسعة XLS (eXtensible Style Language) :

نظام لنماذج الكتابة أو عرض الصفحات يستخدم مع لغة التهيئة والترميز الموسعة XML، وهي مشتقة من نماذج الكتابة المسندلة CSS.

لغة التهيئة أو الترميز الموسعة XML (eXtensible Markup Language) :

إصدارة مبسطة من اللغة المعيارية الموحدة لتهيئة النصوص أو ترميزها SGML قصد منها استخدامها مع المعلومات المباشرة .

بروتوكول Z39.50 :

بروتوكول يسمح للحاسب الآلي البحث في مجموعات المعلومات على النظم البعيدة، كما يقوم بإنشاء مجموعة من النتائج للقيام بمعالجات أخرى، بالإضافة إلى استرجاع المعلومات، وهذا البروتوكول عادة ما يستخدم مع البيانات الببليوجرافية.